OPTIMIZING THE USE OF SYSTEMS ENGINEERING ON PROPOSALS

by

CLEMENT DOUGLAS SMARTT

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2013

Acknowledgements

I would like to recognize some of the people who inspired me to pursue a doctorate degree, who helped guide me through the process and who supported me in other very important ways while I undertook my research.

Inspiration came from a number of places and a number of people, some of who have PhDs and some who did not. In my work at both L-3 and at the Georgia Tech Research Institute (GTRI) I have been consistently impressed with the research skills and pain tolerance of coworkers with PhDs. Individuals who I've encountered with PhDs who I have found especially inspiring from my L-3 days are Dr. Marvin Eargle, Dr. Mike Grabbe and the late Dr. Rosemary Edwards. The number of inspirational PhDs I have encounter at GTRI are too numerous to mention. I am also inspired by my sister's scholarly pursuits and the encouragement of a personal mentor of mine, Dr. Carla Herren who encouraged me to step outside my comfort zone and more fully develop myself intellectually.

I owe a debt of gratitude to many who have directly helped me earn the degree. First and foremost I wish to thank Dr. Susan Ferreira for several hundreds of hours over the years of mentorship and critical feedback. Other committee members, Dr. Rosenberger and Dr. Corley have also have also invested significant time in me and have given me valuable feedback and helped shape my research. There are many other people who have provided me their professional opinions, reviewed various drafts of my dissertation work products, reviewed journal and conference paper drafts, responded to my survey and validated my model. I am in debt to all of those people. I also owe GTRI for offering me the financial support in various forms and the flexibility in my job responsibilities to allow me to accomplish this.

Most importantly, I want to thank the people who are very close to me for making this possible. This includes my wife, Kelly and my daughter Olivia, who have spent many weekends and holidays without me. Kelly, your patience and faith in me will be appreciated for the rest of my life. Other people who made this possible are my parents, Herb and Pat Smartt, who have provided all kinds of help and support throughout this process. Thank you.

November 1, 2013

Abstract

OPTIMIZING THE USE OF SYSTEMS ENGINEERING ON PROPOSALS

Clement Douglas Smartt, PhD

The University of Texas at Arlington, 2013

Supervising Professor: Susan Ferreira

Organizations whose primary business is executing contracts must be able to capture contracts to survive. When the contracts involve engineering complex systems, systems engineering often plays a significant role in the proposal process, sometimes leading the technical effort. This research seeks to find an optimal use of systems engineering in proposal management to maximize the probability that a supplier organization will be awarded contracts.

A number of systems engineering related factors that can potentially be used to predict contract awards are identified that pertain to the organization, the skill levels of employees, the competitive environment, the proposal project, the contract, and the relationship with the customer. A survey was conducted to gather information related to these factors as well as contract award status for recent proposal efforts. An analysis of the survey results indicates that suppliers seeking to be awarded new contracts should: (1) keep their existing customers very satisfied with the contract work already captured, (2) invest adequate resources in systems engineering labor to understand the requirements and define a solution in support of the proposal, and (3) maintain an adequate number of face-to-face contacts with the customer during the proposal process.

A modeling framework was developed and validated to help decision makers determine an optimal use of systems engineering on their proposals. The framework

allows users to maximize the probability of a contract award given constraints, such as budget and employee availability, by strategically allocating resources to key systems engineering activities and employee with various skill levels. Organizations that engineer complex systems can use the findings of the survey analysis and the modeling framework to improve the chances of survival for their organizations.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction

1.1 Overview

Organizations that thrive on contract work must pursue and successfully capture contracts to survive [Gann and Salter, 1998]. Such organizations often invest limited overhead when preparing proposals. If an organization consistently expends resources preparing proposals, but is not awarded sufficient contract revenue to survive, they can exhaust their resources and go out of business. Many sub-organizations within the organization generally participate in and contribute to the proposal process. When complex products or services are being proposed, systems engineering is a key contributor.

Systems engineering is defined as “an interdisciplinary approach and means to enable the realization of successful systems” [INCOSE, 2004]. Systems engineering encompasses the entire product lifecycle, from product conceptualization through disposal. This includes the portion of the lifecycle where opportunities for future work are identified and proposals are prepared. In practice, systems engineering often represents engineering in the early, conceptual stages of an effort and leads the technical effort preparing proposals. Therefore, it is important to understand how systems engineering should be used when preparing proposals. Research has recently been underway to explore how systems engineering can be used to structure the proposal preparation process [Philbin, 2008], to identify stakeholders and metrics for the proposal process [Smartt and Ferreira, 2012c], to define a proposal process flow from a systems engineering perspective [Smartt and Ferreira, 2012c], and to identify a number of proposal related research challenges and opportunities for systems engineering [Smartt and Ferreira, 2011]. This dissertation pursues one of the identified research opportunities in Smartt and Ferreira [2011]: determine an optimal use of systems engineering in the proposal process.

1.2 Statement of Problem

There are many ways systems engineering can contribute to the optimization of the proposal process. Systems engineering can be used to define the process that an organization uses to respond to proposal opportunities [Philbin, 2008]. Systems engineering uses many business and technical processes [ANSI, 1999; ISO 15288, 2008; FAA, 2006] and can recommend that a set of these processes be regularly followed when an organization executes a proposal effort. Systems engineering can also define tailoring guidelines to ensure that the amount of rigor when applying the processes is adequate while satisfying schedule and budget constraints.

Systems engineering also identifies stakeholders and seeks to understand their perspectives. There are a number of stakeholders for a proposal effort, and each has a unique perspective and set of objectives to be optimized [Smartt and Ferreira, 2012c]. Of interest are the stakeholders who must make systems engineering related resource allocation decisions on proposals. One common objective that these decision-making stakeholders all share is the desire to prepare proposals that lead to contract awards.

There are multiple stages of the proposal process where systems engineering can be applied. For example, in the very early stages of the proposal process where organizations are selecting the subset of proposal opportunities that they may wish to pursue, systems engineering provides insight into technical competencies (e.g., the organization’s technical expertise related to particular system types). By providing this insight, systems engineering is helping the organization align proposal opportunities with organizational competencies.

Another stage in the proposal effort often led by systems engineering is defining the proposed offering. Once an organization has made at least a preliminary decision to pursue a particular proposal opportunity, the next step is to decide what will be proposed. This step includes analyzing the acquirer and other stakeholder needs to define and validate source requirements, defining and validating system technical requirements, defining and validating architecture alternatives (including identifying candidate technologies for insertion), estimating costs, and performing decision analysis (e.g., trade studies) for final solution alternatives. This dissertation focuses on how to allocate systems engineering resources in defining the proposed offering to maximize the probability of being awarded a contract within proposal project constraints.

1.3 Research Question

The primary research question is “When defining the proposed offering in a proposal, what is the optimal use of systems engineering to maximize the probability of being awarded a contract?” The phrase "use of systems engineering" includes multiple considerations. These include how much systems engineering to perform and what are the necessary competencies, experience, education, communications skills and interpersonal skills of individuals performing systems engineering related activities on a proposal. This dissertation focuses on defining a framework for selecting an optimal level of investment for systems engineering labor when defining the proposed offering. The framework can provide guidance as to how to distribute that investment across multiple proposal activities and contributors with different skill levels.

1.4 Research Hypotheses

Major hypotheses are:

1. Investing in a higher number of systems engineering labor hours to support the proposal process will result in a higher probability of contract award than investing in a lower number of systems engineering labor hours.

2. Executing proposal efforts with particular factors at specific levels (e.g., more sophisticated systems engineering process maturity level, greater level of customer satisfaction on past contracts, a higher number of interactions with the customer prior to the proposal submission, elevated competitive rank relative to the competition) will positively impact the probability of contract award.

1.5 Research Objectives

The primary objectives of this research are to:

1. Support decision makers so that they may optimally allocate resources for systems engineering when defining the proposed offering for a particular proposal. Specifically, the research seeks to provide decision makers with a framework to generate recommendations as to how to allocate systems engineering labor between activities and skill levels that conform to proposal project budget and employee availability constraints.

2. Provide general guidance for decision makers related to actions that they may take to increase their organization's level of success with being awarded contracts by analyzing results from a survey analysis.

1.6 Research Scope

After considering numerous perspectives [Smartt and Ferreira, 2012c], this dissertation focuses at the project level from the perspective of a supplier of complex systems. A supplier "provides a product or a group of products to an acquirer" [ANSI, 1999]. The primary stakeholder perspective of interest is someone who must make decisions relating to allocating systems engineering resources for a particular proposal effort. Such stakeholders may include project managers, systems engineering functional managers, lead engineers on proposal efforts or lead systems engineers on proposal efforts. The primary objective explored is to maximize the probability of being awarded a contract.

This research focuses on systems engineering's role in defining the proposed offering. The proposed offering includes an analysis of the requirements and the proposed technical solution. The timeframe of interest within the system's lifecycle is from the point where an initial decision is made to propose a solution (to a perceived or documented customer need) until the proposed offering is sufficiently determined to propose the solution to a customer.

This research focuses on proposals where a customized solution must be developed. Large systems integration projects that include numerous non-development items (e.g. customer off-the-shelf system components, government furnished equipment) are within the scope of this research. Proposal efforts of interest include new systems and enhancements to existing systems. Table 1.1 provides an overview of the scope of this research.

Table 1.1 Scope of Research

Dimension	In Scope
General perspective	System supplier
Level of focus	Project
Stakeholder perspectives	Project manager Functional manager responsible for systems engineering Lead engineer on proposal effort Lead systems engineer on proposal effort
Objective	Maximize probability of contract award
Focus within proposal process	Defining the proposed offering (i.e., requirements and technical solution)

1.7 Research Contributions

This research provides two major contributions:

1.7.1 Factor Relationships Study

A study of relationships between various systems engineering related factors (both technical and management) and the probability of contract award. Contributions include identifying a set of factors that may potentially lead to contract awards and using regression analysis to determine which of those factors statistically correlate with proposal success. This statistical analysis includes an assessment of each factor individually and of multiple factors considered simultaneously.

1.7.2 Optimization Modeling Framework for Using Systems Engineering on Proposals

This research defines a framework called the Systems Engineering Proposal Optimization Modeling Framework (SEPOMF). The SEPOMF helps guide the development of decision support systems to optimize the use of systems engineering on proposals based upon an organization's internal historical proposal data. A decision support system (DSS) is an information system that helps and supports people in a decision making process [Power, 2002]. The SEPOMF prescribes how to generate an objective function that is based on the historical data set available, how to mathematically describe project constraints and how to interpret and solve resource allocation optimization problems. Each DSS developed using the SEPOMF has the potential to help decision makers identify an optimal use of systems engineering on proposal efforts to maximize the probability of being awarded a given contract. More specifically, the solution from each DSS helps decision makers determine how many labor hours to devote to systems engineering efforts in defining the proposed offering. In some cases, the DSS will advise how to optimally distribute those labor hours across a number of activities and mix of skill levels of contributors.

1.8 Organization

Chapter 1 provides a summary overview of the technical areas of focus, an overview of the primary research question, an overview of the primary hypotheses, the research objectives, a definition of the scope of the research, the research contributions, and an overview for how the remainder of the document is to be organized.

Chapter 2 provides background information used as a foundation for research. This section discusses key findings from literature from a number of subjects closely related to the subject matter of the dissertation including systems engineering, proposal management, software engineering, economics, finance, game theory, and business strategy. In addition, Chapter 2 also provides relevant background information pertaining to the research methods employed such as the science of conducting a survey, regression analysis and optimization.

Chapter 3 provides additional detail and clarification related to the research question and hypotheses. Chapter 3 also defines the research design, focusing on methods employed to evaluate the hypotheses and answer the research question. Threats to validity and limitations of the research design are also discussed.

Chapter 4 presents the results of the research. This includes a summary of the survey results and a discussion about the SEPOMF. Both the analysis of the survey results and the observations attained while developing the SEPOMF are used to evaluate the hypotheses and address the research question.

Chapter 5 discusses the impact of this research. Findings and recommendations for decision makers using systems engineering on proposals are discussed. Chapter 5 provides an overview of the research contributions and how they extend existing literature. Future work to employ systems engineering to achieve even greater success with proposal efforts is discussed as well as other related future research.

Appendix A provides the survey instrument for the *Survey to Assess the Relationships Between Systems Engineering Factors and Proposal Success* including all of the explanatory information and questions presented to survey respondents.

Appendix B provides an overview of the survey verification process. This is the process that the survey questions and survey instrument underwent in order to ensure a high probability that the survey was error free and technically successful.

Appendix C provides an overview of the survey validation process. The survey validation process is the process that the survey questions and survey instrument underwent to ensure that the set of questions elicits all of the information necessary to support the research questions and needs of decision support systems, that each question is clearly stated in unambiguous terms, that the explanatory information provides the right level of detail for guidance, and that respondents are not subjected to undue risks by participating.

Appendix D provides a detailed survey data analysis. This includes a discussion of how the survey was implemented and promoted, the criteria used to assess the validity of survey responses, the demographics of the respondents and organizations, the characteristics of the projects analyzed, and a detailed analysis of the relationships between various systems engineering related factors and contract award status.

Appendix E provides details related to the definition of the SEPOMF. This includes general directions for applying the SEPOMF and detailed architecture views of the modeling concept.

Appendix F provides examples where different variations of decision support systems are applied to specific proposal project constraints to find optimal allocations of systems engineering labor.

Appendix G describes the validation process of the SEPOMF. This includes the derivation of questions for a validation package distributed to several subject matter experts in the use of systems engineering on proposals, the inputs of those experts and an analysis of the inputs of those experts.

1.9 Intended Audience for Dissertation

This dissertation is written for audiences with significant depth of understanding of analytical models. It is recommended that if an organization attempts to leverage this research to develop such a DSS, in addition to professionals with a depth of understanding in systems engineering and proposal management, they also include at least one person with technical depth in the areas of regression analysis and optimization.

Chapter 2

Background Information

This section presents information that provides background and context for the research presented. This background information is partitioned into two sections: problem domain and research methods. The problem domain section provides background and context related to the research questions and findings. The research methods section provides background and references related to the particular methodologies that are applied to find answers to the research questions.

2.1 Problem Domain

The focus of the dissertation is the application of systems engineering to proposal management. Therefore, a high-level overview of systems engineering and proposal management is provided. An examination is conducted to provide a detailed breakdown of the proposal process, with a focus on the role of systems engineering. Then, research opportunities for applying systems engineering to proposal management are discussed. The discussion will then focus on one particular research opportunity related to optimally using systems engineering in support of proposal management. Because the use of systems engineering in proposal management is a new research area, much of the background literature used to define key terms and concepts was led by the author. This author-led research is presented in the background section because the concepts and terms are leveraged to help frame the detailed questions, the detailed hypotheses and the research design in Chapter 3.

2.1.1 Systems Engineering Overview

This section provides a high-level overview of systems engineering, focusing on the definition of systems engineering, the benefits attained by systems engineering, and commonly referenced systems engineering processes and standards.

2.1.1.1 Definitions and General Purpose of Systems Engineering

While there is no single, universally accepted definition of systems engineering, the following definitions are provided:

“Systems engineering is a discipline that concentrates on the design and application of the whole system as distinct from the parts. It involves looking at a problem in its entirety, taking into account all of the facets and all of the variables and relating the social to the technical aspects” [FAA, 2006; p. 1-1].

“Systems engineering is an iterative process of top-down synthesis, development, and operation of a real-world system that satisfies, in near optimal manner, the full range of requirements of the system” [Eisner, 2002; p. 5].

“Systems engineering is an interdisciplinary approach and means to enable the realization of successful systems.” [INCOSE, 2004; p. 11].

The International Council on Systems Engineering (INCOSE)'s analysis of these definitions emphasizes the following keywords: interdisciplinary, iterative, socio-technical and “wholeness” [INCOSE, 2010]. Organizations apply systems engineering so that technical problems can be considered in their entirety. Systems engineering manages complexity and change and reduces risk [INCOSE, 2010]. Dedicated studies demonstrate the value of applying systems engineering. For example, the Software Engineering Institute (SEI) at Carnegie Mellon University has conducted an extensive quantitative assessment that finds correlations between mature systems engineering practices and a project’s schedule, budget and technical project performance [Elm et al., 2007; Elm 2012].

2.1.1.2 Systems Engineering Processes and Standards

Systems engineering provides a number of technical and business processes that can be applied to engineer a system. In general, professional organizations that conduct systems engineering each possess their own view and decomposition of systems engineering processes. INCOSE's Systems Engineering Handbook [INCOSE, 2010] invokes the International Standard Organization (ISO) 15288 [2008] set of processes. INCOSE also teamed with the Electronics Industries Alliance (EIA) to develop the EIA-632 Processes for Engineering a System [ANSI, 1999]. The United States Federal Aviation Administration (FAA) has published a Systems Engineering Manual [2006] and the National Aeronautics and Space

Administration (NASA) has published its own Systems Engineering Handbook [2007]. The Department of Defense published the Systems Engineering Guide for Systems of Systems [DoD, 2008] to address the unique challenges posed by utilizing a system of different systems, each of which has its own governance and mission. In general, there is considerable content overlap between the processes identified in each set of standards.

2.1.1.3 Empirical Research in Systems Engineering

The empirically based research in systems engineering is somewhat limited in terms of number of studies. Valerdi and Davidz [2008] identify some of the challenges for systems engineering empirical research: (1) the relative immaturity of the field, (2) the lack of appreciation for empirical research, (3) the lack of access to data, and (4) the lack of accepted metrics. Valerdi and Davidz call for the support of professional societies in supporting systems engineering empirical research. This call is increasingly being answered. For example, NDIA is partnering with the Software Engineering Institute at Carnegie Mellon University and sponsoring some very rigorous quantitative research on the effectiveness of systems engineering [Elm et al., 2007; Elm, 2012]. NDIA and the International Council on Systems Engineering (INCOSE) both supported the research reported in this thesis by providing links to the survey.

One area where a significant amount of focus is being invested in empirical research in systems engineering is for estimating the costs of systems engineering activities. COSYSMO [Valerdi, 2005] adapted the parametric cost estimation paradigm originally developed in software engineering [Boehm, 1981] for use in systems engineering. Since then, much work has been done in the field of systems engineering cost estimation [Fortune et al., 2009; Lane, 2009; Valerdi at el., 2007a; Valerdi et al., 2007b; Wang et al., 2009a; Wang et al., 2009b]. Even still, a study finds that as of 2012, the practicing systems engineering community largely lacks confidence in parametric cost estimation to provide “usable” estimates because of the immaturity of the various characteristics of projects and teams that are used to size projects [Miller, 2012].

2.1.2 Proposal Management

This section will discuss (1) the general themes and thrust of the proposal management literature written to address the seller or supplier's perspective and (2) theory with potential applicability to proposal management. To provide additional perspective, the section then presents relevant proposal management literature from the buyer or acquirer perspective.

2.1.2.1 Consultancy Experience Literature

The primary knowledge base of the proposal management literature from the perspective of an organization supplying or selling a product or service is consultancy experience [Philbin, 2008], and the primary focus is on the mechanics of preparing a proposal and the management of the proposal process [Smartt and Ferreira, 2011]. The literature largely targets managers who find themselves responsible for managing a proposal effort [Garrett and Kipke, 2003; Lewis, 2003; Nickson, 2003; Tweedley, 1995; Whitley, 2006]. In general, these sources span the entire process from an organization identifying an opportunity through contract award. Most references address the need to consider what is required to deliver the product or service, but the focus is not on developing the technical solution. Instead, the focus is providing a step-by-step management guide [Smartt and Ferreira, 2011].

2.1.2.2 Empirical Bidding Literature

One practical bid related question that many companies seek an answer to is whether or not another company is likely to submit a bid for a particular contract. Such knowledge has potential implications to important decisions such as how much profit can be added to the cost when setting the price point. There is empirical literature from the construction industry where multivariate binary logistic regression models have been developed to estimate the probability that a particular organization will enter a bid for a particular job [Lowe and Parvar, 2004; Oo et al., 2007]. In these analyses, factors are analyzed for their statistical correlation within an organization's decision to bid. There has also been recent empirical work to use data from actual projects to create a factor-based cost model to support setting a bid price based on project factors [Wang et al., 2012].

2.1.2.3 Bid and Proposal Related Theory

While the proposal management literature focuses on consultancy experience, a variety of theoretical literature exists that may help guide decision making related to proposal management [Smartt and Ferreira, 2011]. This theory includes concepts such as utility theory, cash flows, real options, modern portfolio theory, welfare economics and auction theory.

2.1.2.3.1 *Utility Theory*

One way to analyze proposal opportunities is in terms of cash flows. Money spent preparing proposals can be viewed as cash outflows and profit gleaned from executing contracts that are awarded as a result of successful proposals can be viewed as cash inflows. For opportunities where there is significant time that elapses between the proposal effort and the date at which the contract is complete, the time value of money needs to be considered and out-year cash flows need to be appropriately discounted. However, in order to fully characterize an organization's attitude towards investment in preparing proposals, utility theory can be useful [Smartt and Ferreira, 2011]. Utility theory involves the use of utility functions. A utility function is a theoretical construct that allows the desirability of different outcomes to be expressed based on the value to an individual or organization [von Neumann and Morgenstern, 1944]. Utility theory is useful for analyzing both preferences and risks. To illustrate the explanatory power of a utility function, consider a person with a net worth of $800,000. That person would generally prefer a 50% chance of winning $10 to a 100% chance of being awarded $4 because the expected value of the first option is higher [Smartt and Ferreira, 2011]. However, that same person would likely not bet the entire $800,000 fortune on a 50% chance of receiving $2,000,000, even though the expected value of the riskier option is higher. A utility function is at the heart of what makes insurance attractive to consumers. The expected value of the payout for an insurance policy is always less than the cost. However, consumers of insurance prefer the certain but relatively small economic loss associated with buying the insurance to the uncertain, but potentially catastrophic economic loss that they will suffer if a particular event occurs [de Neufville, 2003].

2.1.2.3.2 Real Options

The term "real option" was introduced by Myers [1984] and is defined as a "right, not an obligation" [de Neufville, 2003]. In other words, someone with a real option may elect to exercise that option, but is not committed to do so. Real options extend the analysis originated to evaluate financial instruments such as stock options to other sorts of decisions. Research and development can be viewed as an example of real options [Myers, 1984]. If an organization elects to invest in research and development, then they are hoping to ultimately obtain the right to enter the market with a product [Myers, 1984]. However, they are by no means obligated to enter the market if conditions are not ripe to earn a profit.

Real options are applicable to analyzing the value of a proposal opportunity because before a proposal is submitted, an organization is investing money to explore an opportunity without a firm commitment. As more information is attained, the estimated value of the proposal opportunity is refined. Real options analysis allows organizations to periodically review and reassess whether continued pursuit of a proposal opportunity remains in the organization’s best interest. Once a proposal is submitted, however, the applicability of the real option analogy is challenged because the organization is often committed to actually deliver the product or service if a contract is awarded [Smartt and Ferreira, 2012c].

2.1.2.3.3 Modern Portfolio Theory

Since proposal opportunities are a type of investment, theory relating to investment management has potential applicability. Investment management theory has historically been applied to analyzing research and development portfolios [Collier, 1968]. One theory commonly applied to assessing stocks and bonds that has potential explanatory value for selecting which proposals to pursue is modern portfolio theory [Smartt and Ferreira, 2011]. A key concept necessary to understand modern portfolio theory is net present value (NPV). A NPV is the estimated value of an investment if all net cash flows (sum of inflows and outflows) are discounted to the present time to reflect the time value of money over some predefined period of time. Markowitz [1952] observed that attempting to maximize NPV without considering risk often favors high risk, high reward options that are in reality not attractive to decision makers. Modern portfolio theory allows decision makers to maximize the NPV within the constraints of a specified risk level. However, this theory has limited applicability to the proposal effort scenario. First, modern portfolio theory assumes earnings are directly proportional to investment. For proposal efforts, there is a minimum investment that must be made to develop a studied response [Smartt and Ferreira, 2011]. Also, there is likely a point of diminishing returns in investing in a proposal effort beyond which few additional benefits are attained [Smartt and Ferreira, 2011]. In addition, modern portfolio theory assumes that potential earnings for a given investment opportunity are identical for all investors. In reality, some organizations may be better positioned to profit from a particular proposal opportunity than others [Smartt and Ferreira, 2011].

2.1.2.3.4 Welfare Economics

Another economic paradigm through which to consider proposal opportunity investments is welfare economics [Arrow, 1954]. Welfare economics describes market transactions in terms of states and preferences. The fundamental concept is that each market participant makes decisions to maximize its overall welfare based upon its current state and its preferences. In the context of proposal efforts, an organization's readiness to compete for a contract and its positioning can be viewed as its "state" and the desirability of the work is its "preference" [Smartt and Ferreira, 2011]. Whether an organization elects to submit a proposal for a particular contract or how many labor hours an organization is willing to invest in a proposal effort is a function of states and preferences. One of the attractions of welfare economics is that everyone benefits from each transaction. However, in the competitive pursuit for contracts, not everyone wins. The possibility exists that an organization may invest considerable amounts to study a problem and prepare a proposal, yet not be awarded a contract. In such a case, the organization may have nothing to show for its investments. Therefore, a fully descriptive model for the proposal economics must consider the concepts of winning and losing [Smartt and Ferreira, 2011].

2.1.2.3.5 Auction Theory

Game theory generally considers decision making in an uncertain environment where the actions of other participants is unknown a priori. In game theory, participants can win or lose. One type of game that has applicability to proposal management is auction theory [Smartt and Ferreira, 2011]. Auction theory is a subset of game theory where multiple participants submit a bid and only a subset are awarded prizes. An auction is defined as “a market institution with an explicit set of rules for determining resource allocation and prices on the basis of bids from market participants” [McAfee and McMillan, 1987, p. 701]. Many people equate auctions with a specific type of auction called an English auction where multiple bidders submit bids and the highest bidder obtains the prize. However, English auctions are only one of many types of auctions.

The auction most applicable to the proposal scenario is an All-Pay auction [Smartt and Ferreira, 2011]. In an All-Pay auction, each bidder must pay the amount bid, but only one bidder (often the highest) wins the prize. Variations include cases where multiple bidders win prizes and other cases where bidders have asymmetric information such that the winner may not be the participant with the highest bid, although in most cases, a more substantial investment leads to better chances of winning [Siegel, 2009]. Political campaigns are a real-world example of All-Pay auctions. Each candidate’s party or support base must invest to promote the candidate, but only one candidate (often, but not always, one with considerable financial backing) wins the election. A more potentially applicable example of an All-Pay auction for proposal efforts is a research contest [Smartt and Ferreira, 2011]. Research contests model various organizations investing in research and development with the objective of achieving a technological advancement before the competition [Che and Gale, 2003].

2.1.2.4 Other Analytical Bidding Models

There are other analytic models beyond game theoretic approaches that can be used to analyze bidding. A few are discussed here to provide a sense of the variety of model types. There are analytical models designed to optimize the price point of a bid to maximize the expected profit [Carr et al., 1983; Hosny and Elhakeem, 2012]. There are also models that use logistic regression analysis to predict whether an organization will submit a bid for a particular bid opportunity [Lowe and Parvar, 2010; Oo et al., 2007]. Finally, some models use systems dynamics to analyze how the boom and bust cycles in organizations whose work is funded via submitting proposals is a result of the proposal policy [Bayer and Gann, 2006].

2.1.2.5 Acquisition Management

In general, there is significantly more literature written about acquisition from the perspective of the acquirer than from the perspective of supplier. Much of the acquisition literature is of interest to suppliers because it provides insight into how buyers make decisions. The proposal literature from the supplier perspective recognizes this by addressing how buyers evaluate proposals [Lewis, 2003]. Some of the supplier-oriented proposal literature suggests strategies (e.g., compliance matrices and solution linkage plans) for having a proposal be favorably received by the acquirer [Garrett and Kipke, 2003]. This section provides a brief overview of the acquirer literature with a focus on content that is potentially useful for suppliers as well as acquirers.

2.1.2.5.1 General Guidance

In the United States, the Department of Defense maintains a guidebook for how to acquire systems [DAU, 2012] including the definition of a systematic process for identifying gaps in current capabilities and targeting acquisition towards fulfilling those gaps [DoD, 2011]. In addition, there exists a university with a primary mission of supporting the acquisition process [DAU, 2013].

2.1.2.5.2 Organizational Buying Behavior

An organization supplying products or services may benefit by looking beyond the general acquisition guidance that focuses on the acquisition process and focus on how buyers make decisions. Organizational Buying Behavior (OBB) is a field of marketing devoted to understanding how organizations make purchases [Sheth, 1996]. OBB extends the largely psychology-focused marketing subfield of Consumer Buying Behavior (CBB) to consider the impacts of conflict and group dynamics, both within the buying organization and between the buying organization and potential organizations seeking to sell products or services [Sheth, 1996].

2.1.2.5.3 Acquisition Management Analytical Models

Part of the acquirer-oriented literature presents analytical models designed to understand and improve the acquisition process. In the systems engineering literature, an incremental commitment model has been constructed to help organizations acquiring products or services by using risk assessments as a criterion for commitment decisions [Boehm and Lane, 2007]. A detailed discreet-event model of the Defense Acquisition Guide (DAG) process has also been created to help identify opportunities for streamlining the United States Department of Defense acquisition process [Wirthlin, 2009]. Other models assume a broader perspective. One system dynamics model focuses on disconnects between suppliers and acquirers as a major source of failure and seeks to alleviate such disconnects by having suppliers increase their understanding of the acquirer's perspective [Greer et al., 2006].

2.1.3 Identifying the Role of Systems Engineering in Proposal Management

Systems engineering plays a role in proposal management. Philbin introduced a high-level model for the proposal process focused on systems engineering considerations in proposal management such as requirements capture, architecture, development and evaluation [Philbin, 2008]. Smartt and Ferreira [2012c] extended the process model to provide a more detailed definition of the activities and their relationships. The Smartt and Ferreira model defines a number of activities and provides insight about what sorts of information flows from one activity to the next. The activities include: (1) make a preliminary pursue/no pursue decision, (2) define the proposed offering, (3) set the price point, (4) make a final decision about whether the organization wishes to actually submit the proposal and (5) prepare the final proposal package.

Making a preliminary pursue/no pursue decision is generally a collaborative process between organizational functions. The ultimate decision may be made by senior management, but systems engineering often provides inputs. While marketing or sales may possess the expertise relating to the competition, follow-on opportunities and the market value of an offering [Young, 2000], systems engineering is positioned to provide input related to technical achievability and costs. Costs, along with market value, determine the per unit profitability of an opportunity. For complex systems, substantial technical expertise may be required to estimate even rough order of magnitude cost estimates [Smartt and Ferreira, 2012c]. Considerable literature exists to facilitate making pursue/no pursue decisions for proposal opportunities. In addition to the broader investment management literature such as portfolio selection [Markowitz, 1952], welfare economics [Arrow, 1951; Arrow, 1954], utility theory [von Neumann and Morgenstern, 1944] and real options [de Neufville, 2003; Myers, 1984], there exists literature dedicated to applying analytical techniques specifically to qualifying proposal opportunities. Approaches exist for applying multi-criteria assessment to determining the probability than an organization can ultimately win in a competitive proposal environment [Cagno et al., 2001] and for applying multi-criteria assessment to determine the "value" of various proposal opportunities [Paranka, 1971].

If a decision is made to pursue, then a proposed offering must be defined. The focus of the research in this dissertation pertains to defining that proposed offering, including an estimate for the costs. Table 2.1 provides an overview of the various activities included in defining the proposed offering, including a definition and the expected outputs.

Figure 2.1 provides a graphical depiction of the sequence for defining the proposed offering. While the flow is depicted as completely sequential, there is sometimes the need to return to previous steps and perform some rework. This is taken directly from Smartt and Ferreira [2012c]. Table 2.2 provides a mapping between the sub-activities that are conducted as part of defining the proposed offering and the ISO-15288 technical processes.

Table 2.1 Defining the Proposed Offering- Activities

	Activity	Definition	Outputs
1	Identify source requirements	The source requirements are all requirements and constraints imposed by the customer and other stakeholders.	Source requirements
2	Validate source requirements	Validating source requirements includes searching for conflicts or voids as well as obtaining customer or other stakeholder buy-in.	Validated source requirements
3	Identify system technical requirements	This activity involves analyzing the source requirements (which are often stated in user or operational terms) in order to formulate a set of system-level technical requirements.	System requirements
4	Validate system technical requirements	This involves searching for conflicts and voids within the system technical requirements as well as attaining stakeholder buy-in.	Validated system requirements
5	Define architecture alternatives	This involves defining alternative system architectures at both a logical and physical level. This also includes identifying and analyzing key technologies.	Architecture alternatives
6	Validate architecture alternatives	This involves searching for conflicts and voids within each architecture concept as well as ensuring that all requirements are addressed by each architecture.	Validated architecture alternatives
7	Estimate cost for each alternative	This includes using various approaches to estimate the relevant costs for each alternative.	Technical solution alternatives (including architectures, technologies and costs)
8	Select preferred solution	This involves applying multi-attribute decision analysis (e.g. trade study) to evaluating the various solution concepts and selecting a preferred concept.	Preferred solution information (architectures, technologies and costs)
9	Validate preferred solution	This involves reviewing the preferred solution with stakeholders such as customer representatives or red teams and ensuring that consensus exists that it is truly the best solution achievable considering the circumstances.	Validated preferred solution

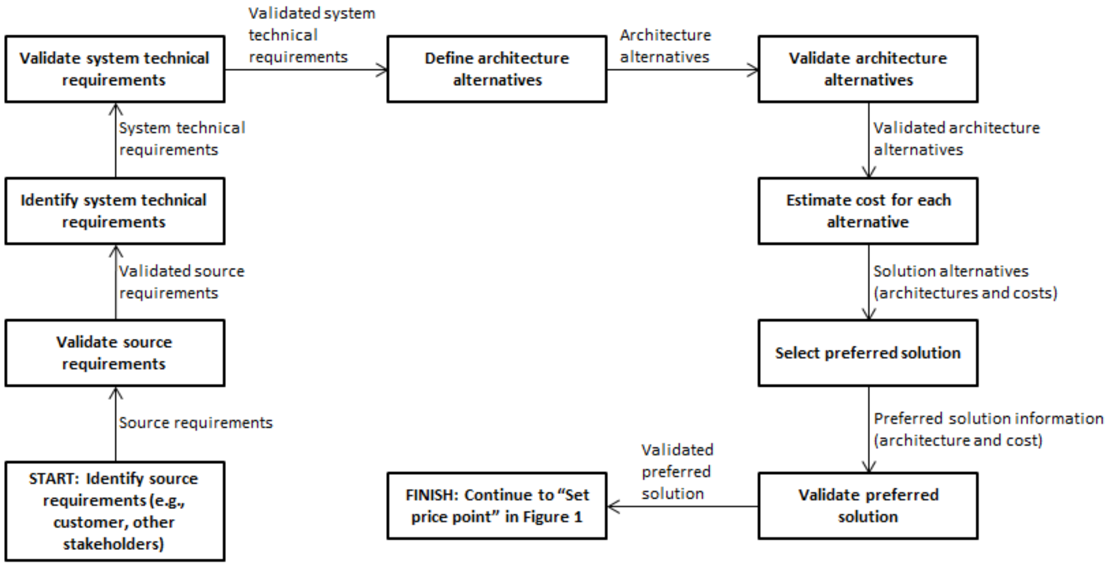


Figure 2.1 Sequence for Defining the Proposed Offering (from [Smartt and Ferreira, 2012c])

Some project processes coincide with all of these key systems engineering activities. For example, decision management, risk management and measurement are done in conjunction with the other activities. Decisions are key throughout all activities. For requirements, an organization must carefully examine their ability to satisfy as well as the cost of satisfying each requirement to determine what subset of requirements that the solution should address. Architecture involves extensive decision making between alternatives. Decision management is the primary process in the decision analysis that includes the final trade studies. The entire process of defining the proposed offering is one of risk reduction and analysis. Requirements are evaluated for risk, architectures and constituent technologies are evaluated for risk, considerable effort is put into understanding the risks of cost estimates, and risk plays a critical role in selecting a preferred solution. Usually, projects will have several key metrics that relate to system performance, cost and schedule. These metrics are tracked throughout the system lifecycle, including during the proposal phase. The values of these metrics are often critical in decision making, potentially including the decision as to whether or not to submit a proposal.

Table 2.2 Mapping of Activities to ISO-15288 Technical Processes

	Source Requirements	System Technical Requirements	Systems Architecture	Cost Estimation	Decision Analysis
Stakeholder Requirements	X				X
Requirements Analysis	X	X			X
Architecture Design			X		X
Implementation			X		X
Verification			X	X	X
Validation	X	X	X	X	X
Decision Management	X	X	X		X
Risk Management	X	X	X	X	X
Measurement	X	X	X	X	X

Considerable literature exists providing guidance for how to set a price point [Agrawal and Ferguson, 2007; Bertolini et al., 2006; Boughton, 1987; Grobler et al., 2008; Kerzner, 2003; Nickson, 2003; Paranka, 1971; Paul and Gutierrez, 2005; Tweedley, 1995]. Some of the literature provides guidance for how to determine an appropriate markup for a defined solution [Bertolini et al., 2006; Kerzner, 2003; Nickson, 2003] while other assumes more of a design-to-cost perspective [Tweedley, 1995]. Systems engineering may provide much of the content that is to be presented in the final proposal package, but the preparation of the final proposal package often is done by graphic arts departments and final wording is set by marketing and approved by senior management.

2.1.4 Strategically Using Systems Engineering in Proposal Management

In order to provide effective guidance related to how to optimally employ systems engineering to support proposal management, it is essential to identify what attributes of systems engineering can be strategically managed to obtain the most favorable outcomes from the proposal process. A starting point is to identify what factors can be managed. A systems engineering strategy framework was recently introduced [Smartt and Ferreira, 2012b] to define the concept of systems engineering strategy. Underlying this framework is a state-based model that conceptualizes applications of systems engineering strategy as moving from a current state to a desired future state. Each state is characterized by a number of state characteristics. These state characteristics define the set of factors that can be modified or are modified by the application of systems engineering. The state characteristics can be organized by the following categories: organization, environment, product, employee, skill level and process [Smartt and Ferreira, 2012a]. The organization characteristics define the organization at whatever level is relevant for making a strategic decision, whether at the whole enterprise level or at the project or team level. These factors include organizational structure and cultural factors [Smartt and Ferreira, 2012a]. The environment characteristics consider customers, competitors, potential partners and regulators [Smartt and Ferreira, 2012a]. The product characteristics describe the particular product being delivered [Smartt and Ferreira, 2012a]. The employee characteristics describe each employee who participates in systems engineering in terms of various characteristics that are applicable to most strategic considerations [Smartt and Ferreira, 2012a]. For each employee and each skill the employee performs, there exists a skill level [Smartt and Ferreira, 2012a]. As systems engineering consists of a number of processes, an important strategic aspect of systems engineering is how processes are applied [Smartt and Ferreira, 2012b].

A number of these characteristics can be managed to realize a desired outcome. The set that can be managed depends upon the context of the strategic decision being supported. This research focuses on defining the proposed offering at the project level. The primary categories of characteristics that are relevant to defining the proposed offering are the product characteristics, the employee and employee skill level characteristics and the process characteristics. The product describes what is being defined. The employee characteristics and employee skill level define who will be participating in that definition and the level of skill of those people. The processes describe how these employees will go about defining the proposed offering. This research will primarily focus on making optimal decisions relating to the use of processes and selecting employees with the optimal skill levels. While product attributes are of essential importance, these vary significantly from one application to the next, and this research seeks to find strategies that are applicable beyond particular products.

The environment and organization must also be considered when defining the proposed offering. However, in many cases, these characteristics may be somewhat determined when a proposed offering is being defined for a particular proposal opportunity. For example, the organization culture and structure is already established and the organization operates at some defined process maturity level. There are exceptions. For example, if the organization is viewed at the project-level, the selection of individual personalities of those tasked with defining the proposed offering may affect the level of harmony between employees. While the environment is also largely determined, there are attributes of the environment that can be changed when defining the proposed offering. In fact, the most fundamentally successful applications of strategy change characteristics in the environment such as creating user base interest in new capability [Smartt and Ferreira, 2012b].

One concern relating to how to use systems engineering on proposals relates to how much to invest in systems engineering on proposals and how to allocate those resources. The various activities described in Table 2.2 provide some investment options for using systems engineering on proposals. While the need for predicting how overall labor will be distributed across various activity profiles has been recognized in systems engineering [Valerdi et al., 2007a], very little analysis has been done to date to understand how to allocate systems engineering effort across lifecycle phases, which are in some ways analogous to the systems engineering activities. Some analysis, however, has been performed in software engineering. The original COCOMO model devotes a chapter to the optimal distribution of projected software engineering labor across life cycle phases and activities for a number of different scenarios

[Boehm, 1981]. More recent studies have been conducted to analyze potential factors that affect the phase distribution as well as "adjustments" to a baseline distribution that can be applied to reflect certain levels for each of these factors [Yang et al., 2008] and to leverage dynamic programming to optimally allocate labor between tasks [Yiftachel, 2006]. However, even in the software development literature, there are just a few papers that address allocation of effort across phases [Yang et al., 2008].

2.1.5 Optimizing the Use of Systems Engineering in Proposal Management

This dissertation addresses the challenge to optimize the use of systems engineering in the proposal process. This section will review existing research that focuses on optimizing the application of systems engineering and optimizing the proposal management process. Then, the topic of optimizing systems engineering in the proposal process will be explored.

2.1.5.1 Optimization in Systems Engineering Literature

In recent years, there is increased focus on determining how to optimally use systems engineering. Such research seeks to rigorously quantify what the optimal amount of systems engineering is for a given challenge [Boehm et al., 2008], how that optimal amount of systems engineering should be adjusted for projects with particular attributes, and how the optimal amount of systems engineering should be distributed across the system development lifecycle stages [Honour, 2011]. Because much of the benefit of systems engineering is risk reduction, there is debate about how effectively the benefits of applying systems engineering can be quantified [Boehm and Sheard, 2010]. One position is that credible, universally accepted estimates for the benefits of applying systems engineering do not exist and will not exist [Sheard and Miller, 2000]. Nonetheless, the movement to optimize the use of systems engineering shows a progression toward a cost/benefit analysis perspective in systems engineering research.

2.1.5.2 Optimization in Proposal Management Literature

One general question that comes to mind relating to preparing a proposal is how many labor hours to invest in the proposal effort in order to obtain the best results. There is some discussion in the proposal management literature addressing this, but it tends to be rule-of-thumb-based [Smartt and Ferreira, 2011]. One example is a recommendation to spend 5% of the proposal earnings for preparing a proposal for "small to medium" jobs, but only about 1.5% for "large" jobs [Lewis, 2003]. Another reference provides a sample cost allocation for a proposal effort [Nickson, 2003]. A third reference provides a curve that suggests an amount to invest in a proposal effort as a function of the projected revenue [Tweedley, 1995]. There seems to be a level of consensus that somewhere between 1% and 5% of the projected revenue should be spent on a proposal effort, but no authors present a rigorous justification for why this is the case.

2.1.5.3 Optimizing the Use of Systems Engineering in Proposal Management: A Stakeholder Perspective

Systems engineering is generally responsible for identifying stakeholders and understanding their requirements. This applies to proposal management as well. Almost all stakeholders or participants in a proposal effort benefit in some way from a successful proposal process. The exact people filling the stakeholder roles vary from project to project as the domain expertise required changes, and in some cases, one individual may simultaneously serve in multiple stakeholder roles. This research elects to focus on stakeholders within the seller or supplier organization. However, each set of stakeholders has a set of objectives they wish to achieve through the proposal process.

2.1.5.3.1 Proposal Effort Stakeholders

This section focuses on the stakeholders within the seller or supplier organization who benefit from a successful proposal process. Table 2.3 lists a set of stakeholder roles and provides definitions of each role. The information in Table 2.3 is an update and extension of the roles discussed by Smartt and Ferreira previously [2012c].

Table 2.3 Proposal Stakeholder Definitions and Roles

Stakeholder Role	Definition
Company Shareholder	A shareholder is part owner of an organization who has a direct stake in the financial success of that organization.
Top Executive	A top executive devises strategies and formulates policies to ensure that organizational goals are being met, and ensures that organizational operations are conducted accordingly. Top executive roles include chief executive officer, chief operating officer, general manager or president [BLS, 2011].
Program Manager	A program manager is the individual with the responsibility for and authority to accomplish program objectives for development, production, and sustainment [DAU, 2013].
Project Manager	"The person appointed to take day-to-day responsibility for management of the project throughout its stages" [Leach, 2000].
Capture Manager	A person responsible for winning a business opportunity [CapturePlanning.com, 2013].
Lead Engineer	The lead engineer coordinates the engineering aspects of a project. This individual interfaces with both project management and the functional managers of the organizations providing engineering expertise to ensure that the right individuals are contributing to the proposal effort, and that they are contributing in the right ways.
Lead Systems Engineer	The lead systems engineer coordinates the systems engineering effort on a project. This individual interfaces with project management, the lead engineer and functional managers of the organization providing systems engineering expertise to ensure that the right individuals are contributing to the systems engineering portion of the proposal effort, and that they are contributing in the right ways. The lead systems engineer is also responsible for ensuring that the technical effort is conducted in an integrated way. This includes establishing and managing the integrated product and process teams. In some organizations, the lead engineer oversees all the technical projects in the organization. When this is the case, the lead systems engineer may be the top technical authority on a project.
Functional Manager of Organization Performing Systems Engineering	In matrix organizations, there are often organizations chartered to perform a particular function. Employees are often managed by these functional organizations and assigned to work particular projects. The managers of these functional organizations generally assign particular individuals to work on particular proposal efforts. The functional management also generally defines the processes by which work is to be executed.
Individual Contributor	This person contributes directly to a proposal effort. There is often more than one.

2.1.5.3.2 Objectives

Each stakeholder has a set of objectives that they may wish to accomplish through the proposal process. Table 2.5 describes the set of potential stakeholder objectives. Each stakeholder seeks to satisfy a subset of these objectives. Table 2.4 and the related discussion will explain the relationship between stakeholders and objectives.

2.1.5.3.3 Mapping of Stakeholders to Objectives

A previous version of Table 2.4 has been presented to relate stakeholders to objectives [Smartt and Ferreira, 2012c]. Table 2.4 updates that previous work based upon further research and the addition of the capture manager role.

Table 2.4 Mapping of Stakeholders to Objectives

	Maximize Value of Real Options	Maximize Value of Net Present Value	Maximize Probability of Winning Proposal	Minimize Cost of Preparing Proposal	Maximize the Capital Position of the Organization	Minimize the Idle Rate of the Staff	Maximize the Likelihood of Repeat Business
Company Shareholder					X		X
Top Executive					X		X
Program Manager	X	X	X	X			X
Project Manager	X	X	X	X			X
Capture Manager			X	X			
Lead Engineer		X	X				X
Lead Systems Engineer		X	X				X
Functional Manager		X	X			X	X
Individual Contributor			X			X	X

Company shareholders and top executives tend to focus on the overall profitability or capital position of the organization and see proposal efforts as a means to secure this profitability. Company shareholders and top executives may also prize lasting relationships with repeat customers.

Program managers tend to focus at the program level. Program managers seek to make their programs as profitable as possible. Project managers tend to focus at the project level. They seek to make their projects as profitable as possible. Both program managers and project managers often assess that profitability through analyzing real options and NPVs. They understand that winning contracts and keeping costs as low as possible is a prerequisite to profitability. They also attempt to foster lasting relationships with customers. The reputation of a program manager or a project manager in many cases may be bolstered by a customer returning.

Capture managers focus on a strict subset of the objectives of a project manager such as winning the contract and minimizing costs.

The engineering leadership on the project (both lead engineer and lead systems engineer), are seeking to win the proposal, but also wish to reduce risk and thus increase the overall profitability of the project. Because of the ubiquitous teaching of NPV in engineering and economics curricula, these individuals may largely view profitability in terms of NPV. The lead engineer and lead systems engineer may also seek repeat business.

Functional managers also care about profitability in terms of NPV and wish to be participants in winning proposal efforts. However, functional managers are likely to make decisions to minimize the idle rate of their staff. Functional managers may view winning repeat business as an enabler to keeping staff productively assigned to projects.

Individual contributors are also likely to make decisions with the goal of securing future work. Individual contributors often want to avoid idleness by winning proposals and later participating in the contract execution. Contributors sometimes build relationships with specific customers and view continuing business between those customers and the organization as job security [Smartt and Ferreira, 2012c].

2.2 Research Methods

This section will provide background related to research methods employed in this research. First, background related to conducting the survey to be described in Chapter 3 will be presented. Then, decision support systems will be discussed. Decision support systems generally involve the use of some analytical models or algorithms that facilitate the decision process. Some relevant types of models used in this research are then discussed.

2.2.1 Survey

The data used by the analytical models in this research is collected by a survey. This section will provide background information relating to planning an effective survey, formulating effective survey questions, effectively administering a survey and correctly analyzing survey results. There is an extensive body of literature related to surveys, from guidance focused on particular activities that must be done as part of survey research to more holistic guidance that applies to coordinating or managing the entire survey process.

2.2.1.1 Determining the Most Effective Type of Survey

Literature exist to provide guidance as to whether to administer a survey via a questionnaire where respondents directly provide answers and there is no direct interaction between a respondent and an interviewer or whether to send live interviewers to directly question potential respondents [Alreck and Settle, 1995; Fink, 1995; Fowler, 1993]. In general, questionnaires are most appropriate (i) where there is no need to interact with the respondent, (ii) where interaction with a respondent may bias the responses, (iii) where a plan exist for how to access qualified respondents, (iv) where respondents have the skill set to record their own responses, (v) where a large amount of data is needed from each respondent, (vi) where respondents are geographically dispersed over a large area, and (vii) where respondents may feel threatened or embarrassed speaking with another human [Alreck and Settle, 1995]. On the other hand, surveys conducted by interviewers are most appropriate for cases where two-way interaction is needed, where respondents are unlikely to have the skill set to record their own data, and where respondents are geographically concentrated [Alreck and Settle, 1995].

Table 2.5 Proposal Management Objectives

Level	Objective	Definition	Advantages	Disadvantages
Project-Level Objectives	Maximize Value of Real Option	Real options were introduced by Myers [1984] and are defined as a "right, not an obligation" [de Neufville, 2003]. Generally, real options, like traditional stock options, can be assigned monetary values. Decisions are made to maximize monetary value.	Real options are flexible in that they support evaluation of an opportunity where no commitment has been made (like determining what requests for proposals to respond to) [Smartt and Ferreira, 2012c].	Once a proposal has been submitted and the customer has selected an organization to perform the contract, that organization may then be committed to executing the contract and the concept of an option is no longer applicable [Smartt and Ferreira, 2012c].
	Maximize Net Present Value (NPV)	An NPV is the estimated value of an investment if all cash inflows and outflows are discounted to reflect the time value of money.	NPV is well understood and taught in many finance-related courses for business and engineering students.	NPV is sensitive to unknowns such as discount rates and cash flows.
	Maximize Probability of Winning Proposal	The probability of winning the proposal is the probability that a contract will be awarded as a result of the proposal submission.	Being awarded contracts is necessary to the survival of organizations that thrive on executing contracts.	Being awarded contracts is not sufficient to ensure that organizations prosper. Organizations must also execute contracts profitably.
	Minimize Cost of Preparing Proposal	The cost of preparing the proposal includes all costs for labor and materials necessary to prepare and submit a proposal.	Costs are generally easy to track.	Looking primarily to minimize costs may lead to inadequate systems engineering rigor [Smartt and Ferreira, 2012c].
Organization-Wide	Maximize the Capital Position of the	The capital position of an organization is its overall wealth, with specific focus on	Making decisions with the objective of maximizing the capital position of the	Understanding the net organizational impacts of choices can be

Table 2.5—*Continued*

	Organization	available cash.	organization incentivizes global optimization as opposed to sub-optimization at the project level.	challenging. Also, finding a specific metric that does not reward sub-optimization of some type is challenging [Smartt and Ferreira, 2012c].
	Minimize the Idle Rate of the Staff	Staff is idle when they are "not occupied or employed" [Merriam-Webster, 2013].	Generally, employees must be paid regardless of whether they are contributing. Therefore, assigning idle employees to pursue future contracts seems an obvious solution to idleness.	Systems dynamics modeling has demonstrated that if an organization routinely assigns employees to proposal pursuit efforts, the magnitude of the fluctuations between an organization being understaffed and overstaffed intensifies [Bayer and Gann, 2006].
	Maximize the Likelihood of Repeat Business	The likelihood of repeat business is the probability that existing customers will return to acquire products or services from the organization in the future.	Repeat business can result in a steady stream of revenue over time with minimal marketing expenses.	Determining exactly how to incentivize particular customers to return may be less than obvious [Smartt and Ferreira, 2012c].

For surveys involving questionnaires, the surveys may be administered via a hard copy physically supplied to respondents or they may be administered via the internet. In general, mail surveys have the advantage that the population of particular respondents can be controlled. On the other hand, anyone with access to an internet survey may respond [SuperSurvey, 2012]. Internet surveys have the advantage that the coding of the responses can be simplified because the responses are collected electronically [SuperSurvey, 2012]. Respondents in internet surveys also provide more complete answers to open questions [SuperSurvey, 2012]. However, response rates for internet surveys are even lower than those for mail surveys [Kaplowitz, 2004].

2.2.1.2 Determining the Sample

A survey designer has options for how to select the sample of respondents to reply to the survey. Random sampling is one of the simplest rigorous sampling methods available and has the advantages that an unbiased sample of the population can be obtained without having to design an intricate experiment [Fink, 1995]. In addition, one can use prediction techniques from inferential statistics with high validity [Alreck and Settle, 1995]. One can also use stratified sampling [Fink, 1995] to ensure that all segments of the population are represented in proportion to the size of the segment of interest to the entire population. This can result in high levels of confidence from a relatively small sample as stratified sampling mitigates for non-response biases by ensuring all strata of the target demographic are assessed and results are appropriately weighted by strata [Frankel, 1983]. A less rigorous approach is convenience sampling. Convenience sampling involves selecting respondents from the population that are obtainable or convenient to reach. Because sampling populations reached by convenience sampling are generally not representative of the larger population, it is technically invalid to apply inferential statistics to results obtained from convenience sampling [Alreck and Settle, 1995].

Another question relating to determining the sample is how large does the sample need to be? In other words, how many respondents need to respond in order for the sample to yield reliable results? The answer varies depending on the data that is analyzed and the model used to analyze the survey results. There are generally complex statistical relationships relating confidence level to the required number of samples. Small sampling size is just one of many possible sources of error in a survey. Expending excessive resources and attention to sample size can detract attention from other very real sources of possible error [Fink, 1995; Fowler, 1993].

Alreck and Settle [1995] claim that for surveys designed to gather statistical data on a particular variable of interest, less than 30 respondents will provide too little certainty to be practical. For a sample of 40, the 95% confidence interval is generally 73% of the estimated mean to 127% of the estimated mean. For a sample size of 300, the 95% confidence interval ranges from about 91% of the estimated mean to about 109% of the estimated mean. On the other hand, if the analytical model that a survey is supporting is a logistics regression model, the number of required responses depends upon the number of parameters in the model and the number of survey responses associated with each possible outcome [Harrell et al., 1985; Peduzzi, 1996].

2.2.1.3 Formulating Effective Survey Questions

For a survey to gather the information that it seeks, it is necessary to write questions that effectively elicit the data necessary to fulfill the survey's purpose. There are open questions and closed questions. While open questions allow users to provide their thoughts, responses are very difficult to code, and less information is often gained from open questions than closed questions [Alreck and Settle, 1995]. The following are generally attributes of good questions:

- Good questions focus on the information being sought by the survey [Alreck and Settle, 1995].
- A good question uses common vocabulary that respondents will be familiar with [Alreck and Settle, 1995; Sheatsley, 1983].

- A good question provides adequate wording for respondents to understand exactly what is being asked of them so that they may provide responses that comply with the data needs of the researcher [Fowler, 1993]. Vague or ambiguous words can elicit responses that do not comply with the data needs of the researcher [Fowler, 1993].
- A good question is grammatically correct [Alreck and Settle, 1995].
- A good question is succinctly written [Sheatsley, 1983].
- A good question provides a balanced range of alternatives, and does not present alternatives that favor a particular type of response [Sheatsley, 1983].

On the other hand, there are some common errors made in question writing. The following is a partial list of common errors in survey questions.

- Double-barreled questions – single questions that ask about two things [Alreck and Settle, 1995; Sheatsley, 1983]
- Overlapping alternatives – response alternatives are not mutually exclusive [Sheatsley, 1983]
- Open questions with examples provided – when examples are provided for open-ended questions respondents often list one of the examples, but fail to think of unlisted examples. Therefore, the responses are biased toward the items that have been included in the list of examples [Alreck and Settle, 1995].
- Intentions to act – asking respondents what they would have done in a hypothetical situation provides unreliable results as respondents are generally not reliable when reporting how they would act in hypothetical situations [Sheatsley, 1983].

2.2.1.4 Validating the Survey Questions

Pretesting is an important activity to help filter out bad questions before the survey is administered. The purpose of pretesting is to determine how the survey instrument and related instructions work under conditions analogous to those the actual respondents will face [Fowler, 1993]. Pretesting can range anywhere from approximately half a dozen respondents to pilot tests where over 100 respondents participate. The set of pretest subjects should be representative of the sample population to be studied. The ideal way to conduct a pretest is in person with a group of test respondents [Fowler, 1993] where test respondents can be "debriefed" after their encounter with the survey instrument [Sheatsley, 1983]. This allows for the test subjects to provide useful insight into what elements of the survey instrument are problematic. If this is impractical, testers should be encouraged to provide feedback in addition to filling out the survey.

In addition to pretesting the survey instrument with respondents, it is also important for the researcher to consider the reliability and validity of the survey as a whole. Ideally, a well-designed survey should meet the following reliability criteria:

- Be stable – If the same respondent was asked the same question multiple times, ideally the responses from time to time would be similar [Fink, 1995].
- Be homogenous – The content and structure of the survey instrument should be as homogenous as possible. All questions should relate to the key information being sought and the items in the instrument should be structured and worded as much alike as possible [Fink, 1995].

The following list describes various dimensions of survey validity:

- Content validity – degree to which a question or set of questions measures what it is intended to measure [Fink, 1995]
- Face validity – does the question or survey seem to ask the right questions [Fink, 1995]
- Predictive validity – extent to which a measure estimated from responses to a survey forecast future performance [Fink, 1995]
- Construct validity – items in survey distinguish between respondents who have or do not have the same key characteristics [Fink, 1995]

2.2.1.5 Effectively Administering a Survey

Significant literature provides guidance as to how administer surveys via telephone, mail or face-to-face interviews [Alreck and Settle, 1995; Dillman, 1983; Fink, 1995; Fowler, 1993]. While conducting internet-based survey research is relatively new, many of the principles that relate to effectively administering surveys in more traditional mediums apply. First, respondents need to be informed in some way about the existence of the survey and the need for their participation. Second, a well-selected inducement can motivate a respondent to participate. The best inducements are generally something that cannot be attained for a reasonable price on the open market or have non-trivial financial value [Alreck and Settle, 1995]. If the survey is announced as it is opened, the majority of the respondents will reply to the survey within a few days of it becoming available [Alreck and Settle, 1995]. Generally speaking, researchers should not pre-determine an amount of time a survey should be accepting data. This allows the researcher the flexibility to begin analyzing the data as soon as sufficient samples have been collected but also make follow up efforts if not as many respondents reply as anticipated [Alreck and Settle, 1995].

For electronically-administered surveys, such as internet surveys, it is imperative to thoroughly test the survey instrument prior to the survey collecting data. As the majority of responses arrive within days of the announcement of the survey, realizing there was a technical error associated with the encoding of the survey instrument could seriously compromise the research project. Like for other forms of software, it is advisable that multiple individuals in addition to the coder test the survey instrument. Each of these testers may attempt to perform operations or select options in combinations that the coder is not thinking about when coding the survey. Catching coding bugs through such testing reduces the probability of failure of the electronic survey instrument.

2.2.1.6 Preparing the Data for Survey Analysis

Prior to statistically analyzing the data, it is necessary to prepare the data for statistical analysis. With traditional survey cards (whether the survey was administered by a professional or by the respondent) responses had to be recorded and encoded via some numerical scheme to facilitate data analysis [Fowler, 1993]. Because internet surveys are software, coding schemes can be defined as the survey instrument is being coded and the step of manually entering the survey data can be bypassed. In either case, a coding scheme must be defined at some point. Regardless of the survey method employed, the survey results must be analyzed for accuracy, consistency and completeness prior to beginning the statistical analysis [Fowler, 1993].

2.2.1.7 Analysis of Survey Results

Usually the goal of research employing surveys is to determine if a factor has correlation with another factor. The survey questions serve to measure, rank or quantify the factors in some way. Because of this, it is common to perform statistical factor analysis and tests of statistical significance related to factor levels of some variables with respect to other variables.

2.2.1.8 More Holistic Guidance to the Survey Process

There is survey research focused on considerations above and beyond the particular activities that are part of survey research. These include how to organize a team of researchers [Iarossi, 2006], how to manage resources for a survey [Iarossi, 2006], how to avoid common management mistakes in survey research [Kuhn, 2012], and how to leverage research design philosophy to the benefit of survey research [Trochim, 2012]. Smartt and Ferreira [2013] discuss how to apply systems engineering processes to survey research.

2.2.2 Decision Support Systems

A DSS helps people make decisions [Power, 2002]. A DSS can allow decision making to be conducted (1) in more-productive ways, (2) with greater agility and alertness to the unexpected, (3) more innovatively, (4) more reputably (e.g., greater accuracy, greater traceability between decisions and inputs), and (5) with higher satisfaction by stakeholders (e.g., decision participants, decision sponsors, decision implementers) [Holsapple, 2008].

2.2.2.1 Application

Decision support systems are used to facilitate decision making for a number of different types of problems in a number of different problem domains. Example problems for which a DSS has been applied include selecting a target demographic for marketing efforts for electronic components to be used in land line telephone infrastructure [Gensch et al., 1990], determining an optimal set of security measures to implement to protect information security systems used for all industries [Strauss and Stummer, 2002], and determining an the optimal profit markup for proposals submitted in the construction industry [Marzouk and Moselhi, 2003]. A DSS helps decision makers make sound choices in the context of semi-structured problems [Keen and Scott Morton, 1978]. Semi-structured problems involve both known and unknown components to the decision making process [Gorry and Scott Morton, 1971]. Decision support systems are most helpful for facilitating decision making where the problem has sufficient structure for a computer and analytic aids to be of value, but where the judgment and insight of a decision maker is required [Keen and Scott Morton, 1971].

2.2.2.2 Varieties of Decision Support Systems

Multiple types of decision support systems exist. These can be categorized along several dimensions. Three such dimensions are: (1) knowledge management techniques, (2) scope, and (3) result interpretation.

Decision support systems vary in their general knowledge management (KM) techniques. While there exist numerous definitions for KM [Doelling and Ferreira, 2010], one that appears particularly applicable is "an explicit, systematic approach for creating, accessing, validating and applying knowledge needed to accomplish goals and objectives" [Carnes, 2002]. DSS knowledge management approaches include: (1) emphasis on the use of document management, which involves the representation and processing of text or hypertext, (2) emphasis on processing information from a data store or database, and (3) emphasis on reasoning and decision making [Burnstein and Holsapple, 2008].

Decision support systems vary in scope as well. A DSS can be used to facilitate one person's decision making, the decision making of an entire organization, or the decision making across multiple organizations. An example of a personal DSS would be an "information dashboard" that keeps an executive apprised of current status related to events that are in the executive's unique purview. An example of an organization-wide DSS is one designed to provide an enterprise's managers with business intelligence of interest to multiple participants in the organization. An example of a multi-organizational DSS would be one that mines information related to the competitive landscape in which an organization operates [Burnstein and Holsapple, 2008].

Decision support systems also differ in their emphasis on reasoning and decision making. This is best illustrated by contrasting expert versus advisory systems. Some decision support systems are classified as expert systems. Expert systems provide definitive solutions to well-defined problems where the steps are well-understood. Advisory systems, on the other hand, are used to provide recommendations to a human user who ultimately is responsible for making the correct decisions [Beemer and Gregg, 2008]. An example of such a system is a graph-based approach to differentiate cancerous from non-cancerous growths in brain tumors [Demir et al., 2005]. Ultimately, the physician is responsible for interpreting the output of the DSS and making decisions related to treatment. The physician may be aware of unique, extenuating circumstances that may cause the results of the DSS to be incorrect, and require that expert judgment and experience override [Beemer and Gregg, 2008].

2.2.2.3 Components and Architectures of a DSS

Like all types of systems, a DSS is divided into architectural components. Consensus in the DSS literature has yet to be reached regarding what defines the top-level architectural components of a DSS. One view defines the components as user interface, model management and data management subsystems. The user interface facilitates the input/output between the system and the user. The model management subsystem contains the models and provides the analytical capabilities for the DSS, and the data management system contains the data, meta-data and data access methods [Turban et al., 2007].

Another view defines the components as language system, presentation system, knowledge system and problem-processing system. The language system consists of all information that the DSS can accept. The presentation system consists of all information that the DSS can supply to the user. The knowledge system consists of all knowledge that the DSS has stored and retained, and the problem-processing system is the DSS's software engine. The problem-processing engine's primary function is to recognize and solve problems [Holsapple, 2008].

Yet a third view defines the components as the knowledge base, the inference engine, the user interface and the explanation facility. The knowledge base of an expert system represents, formalizes and contains the necessary knowledge such that the inference engine can make a decision. The inference engine organizes and controls the processes or algorithms necessary to solve the problem. The user interface allows for the user to input information and receive feedback from the inference engine. An explanation facility provides transparency into the reasoning and information used to arrive at a recommended action [Beemer and Gregg, 2008].

2.2.2.4 Success Criteria and Validation

When a DSS is successful, it can have significant benefits to those who use it. For cases where a DSS is designed specifically to facilitate decision making within an organization, a successful DSS has become an integral part of an organization’s core business strategy and has facilitated sustained competitive advantage [Gench et al., 1990]. Therefore, it is of interest to understand what constitutes success for a DSS. Arnott and Dodson [2008] identify the following critical success factors for a DSS: (1) committed and informed executive sponsor, (2) widespread management support, (3) appropriate team skills, (4) appropriate technology, (5) adequate resources, (6) effective data management, (7) clear link to business objectives, (8) well defined information and system requirements, (9) evolutionary development, and (10) management of project scope. However, for prototypical decision support systems such as the one developed in support of this dissertation, a successful DSS is a DSS that supports a decision that is important and provides sensible recommendations that are based upon analysis of appropriate rigor.

Richardson and Pugh [1981] describes types of validation for system dynamics models. Many of Richardson and Pugh's model tests, however, can be more broadly applied. Table 2.6 presents an overview of a subset of the model tests. Some of the model tests in Richardson and Pugh [1981] have been generalized to apply more broadly than for just systems dynamics models. For a more detailed discussion of these model tests, see Appendix G.

Table 2.6 Validation Model Tests

Activity Type	Structure	Behavior
Consistency	Dimensional consistency	Parameter (in)sensitivity
	Extreme condition test in equations	Structural (in)sensitivity
	Boundary structural adequacy tests	
Suitability	Face validity	Comparison against case studies
	Parameter validity	Surprise behavior
		Statistical tests (where possible)
Model Utility and Effectiveness	Appropriateness of Structure	Counterintuitive behavior
		Generation of insights

2.2.2.5 Decision Support Systems that Consider Multiple Objectives

Research relating to optimizing the use of systems engineering on proposals may eventually seek to address multiple objectives. There exist many different classes of mathematical approaches for solving multiple objective functions. These include (1) aggregation of multiple objective functions into a unique function defining a complete preference order also known as multiattribute utility theory (MAUT) [Olson, 2008; Roy 1971], (2) progressive definition of preferences together with exploration of the feasible set [Roy, 1971], and (3) Analytical Hierarchy Process (AHP) [Olson, 2008].

Multiattribute utility theory involves defining a single formula that allows for alternatives to be evaluated against a number of criteria [Olson, 2008]. Weights are applied to each.

Traditional tradeoff analyses or trade studies in systems engineering [INCOSE, 2010] are examples of MAUT. The advantages to MAUT are that it is conceptually easy to understand and used frequently to facilitate decision making. The disadvantage is that in many cases, MAUT fails to correctly discriminate between alternatives. It may be that all alternatives are acceptable with respect to the most heavily weighted attribute, even though some score better than others. However, one alternative may perform substantially better for a less heavily-weighted criteria than the alternatives. In such a case, the discriminating power may lie in the less heavily weighted alternative [Roy, 1971].

An alternative approach is to explore the feasible set of alternatives by comparing one pair of alternatives at a time [Roy, 1971]. This category of approach is sometimes referred to an “Outranking Method” [Olson, 2008]. Only efficient alternatives are explored. Efficient alternatives are defined as alternatives such that there exist no other alternatives in the feasible set that perform better with respect to one attribute [Roy, 1971]. Essentially, the efficient alternatives are compared one pair at a time until a solution is found that the decision maker prefers to all other alternatives in a one-by-one comparison. More automated metrics are defined to accomplish comparisons when the set of efficient alternatives exceeds what a human can reasonably sort through on a pair-wise basis [Roy, 1971].

The AHP provides a structured approach to comparing alternatives by defining weights for various alternatives at different levels [Taha, 2007] that are determined by using a ratio scale [Forman and Gass, 2001]. AHP has been used for a number of resource allocation problems, including problems for how to fund various research and development projects [Forman and Gass, 2001]. One controversial aspect of AHP is that in its original form, AHP allows rank-reversal. In other words, the introduction of an alternative or elimination of an alternative can cause the remaining set of alternatives to change preference rank. While this may appear problematic, it may be somewhat representative of how people actually make decisions [Forman and Gass, 2001]. There are many decision support systems based on the AHP

[Olson, 2008] that support a variety of decisions, including one that uses the AHP to recommend markups for proposals [Marzouk and Moselhi, 2003].

2.2.3 Relevant Models

This section describes the analytical models and concepts that are appropriate candidates for application in the DSS developed for this research. Three types of models are reviewed: use case models, regression models and optimization models. Regression models are used to predict a quantity of interest based upon a set of values for predictor variables. Optimization models are used to maximize or minimize some quantity of interest given a set of constraints and a value or cost function.

2.2.3.1 Use Case Models

Use cases are a means to specify how a system is to be used [OMG, 2011]. They give a clear description of what a system does or should do [Eriksson and Penker, 1998]. Use cases are generally defined in the context of actors. Actors represent what interacts with the system [Cockburn, 2001], and can be human users or other systems [OMG, 2011]. Use cases and actors are graphically related in a use case diagram. A use case diagram depicts the relationships between actors and use cases. Specifically, a use case diagram displays which actors are associated with which use cases. Formalizing a problem with use case modeling can reveal the extent of the users and functionality associated with a system. Use case diagrams can also be helpful for supporting the definition of functional requirements.

In the context of a decision support system or other type of model, use cases, actors and use case diagrams are helpful for communicating the purpose, audience and scope of the model. Use case modeling helps verify that a model will answer questions that are important to stakeholders, and that the model will support the right set of stakeholders.

2.2.3.2 Regression

Regression analysis provides a way to use actual data to represent a dependent variable as a function of one or multiple independent variables. Regression functions can be used to predict the value of the dependent variable based on the value of one or multiple independent variables. When using regression analysis, it is important to select the type of regression model that is appropriate for the problem at hand. In this research, regression analysis is used to express the probability of an event occurring (i.e., contract award) as a function of various factors relating to the use of systems engineering on proposals.

Probability values must be between 0 and 1. Traditional linear regression models have a domain and range that is the set of real numbers. While only a subset of the domain and range may be meaningful in the context of a particular problem, traditional linear regression models are not inherently restricted to produce responses in a specified range. Therefore, traditional regression models are not appropriate for predicting probability of contract award. Any prediction approach used to predict probability must have a range restricted to [0,1]. Another challenge inherent in the probability of contract award prediction model is that the input data all has a response value of 0 or 1. In other words, for each proposal project, the project either resulted in an award or did not result in an award. Logistic regression models are appropriate for predicting probability of an event based upon data with a binary response [Cook et al., 2001]. Examples of types of problems for which logistic regression is appropriate include predicting the probability that a family will buy a new car as a function of income [Agresti, 1996], and predicting mortality statistics [Perks, 1932]. While computing the parameters of a logistic regression model requires computationally demanding approaches such as maximum likelihood methods or Firth's method [Firth, 1993], off-the-shelf tools such as SAS [SAS, 2013], SPSS [IBM, 2013], logXact [Cytel, 2013] and Matlab's statistics toolboxes [Mathworks, 2013] have functions that automatically estimate the regression parameters based upon particular data.

One major challenge is regression analysis of survey data is how to account for missing data. Missing data is an issue in survey data analysis because respondents often elect not to respond to particular questions. From a mathematical perspective, something must be done about missing data in order for the calculations of the regression parameters to be performed. Table 2.7 provides a high-level overview for some of the general approaches that can be used to address missing data along with the advantages and disadvantages of each approach.

The seemingly obvious approach of “throw out all responses where there is missing data” may in fact lead to incomplete or invalid findings. There can be systematic trends related to respondents who elect not to respond to a particular question. For example, in a study on Optimism and Fundamentalism [Sethi and Seligman, 1993], according to Howell [2013], analysis was done to determine whether or not the level of fundamentalism in a subject’s religious worldview was a predictor of that subject’s overall level of optimism. Subjects who elected not to provide information about their religion were more optimistic than subjects that fell into any of the stated categories (i.e., fundamentalists, moderates, liberals). In this case, non-response was a stronger indicator of optimism than any of the provided options.

2.2.3.3 Optimization

Optimization helps decision makers responsible for proposal success make the best decisions in their situations. Optimization is achieved by minimizing or maximizing a function, usually subject to some equality or inequality constraints [Nemhauser and Rinnooy Kan, 1998]. This function is called an objective function. Constraints define the space of possible solutions to the optimization problem. A solution is feasible if it satisfies the constraints [Taha, 2007]. The set of possible feasible solutions is referred to as the feasible region. The variables that are under the control of the decision maker are called “decision variables”. Decision variables are “a quantity that a decision maker controls” [Optek, 2013].

A value for a function is optimal if it satisfies all of the constraints to a problem and if it yields the best (minimum or maximum) value of an objective function [Taha, 2007]. A number of different approaches are used to determine optimal values. These include calculus-based approaches where one searches for domain values for which the derivative or gradient of the objective function is zero. These also include approaches from operations research such as linear and non-linear programming. These approaches provide methodical ways to locate an optimal value for a function in a feasible region.

Table 2.7 Overview of Approaches to Address Missing Data

Approach	Description	Pros	Cons
List-wise deletion	Simply delete the entire entry from the model building process if ANY of the necessary parameters are missing [Howell, 2013].	1.) Conceptually simple. 2.) Seems like a conservative and highly defensible approach to people without expertise in missing data. 3.) Easy to analyze: does not mix observed data with imputed data.	1.) May result in throwing away most of the survey responses. This leads to lower confidence and less power in results. 2.) Sometimes no response is a statistically significant predictor [Sethi and Seligman, 1993; Howell, 2013]. Just deleting data throws this information away.
Non response category	Add a level for categorical variables that corresponds to "no response was given" [Howell, 2013].	1.) Captures the information inherent in people not responding to questions. 2.) Does not mix observed data and imputed data.	1.) Only sensible for categorical predictors. 2.) Adds an additional level (and hence additional binary variable) for each categorical variable. Drives up variance and drives down confidence levels.
Imputation	Input data for the entry that is somehow derived from other data. This may be taking a populated value from a near-neighbor, randomly picking a value from "like" entries, or a more sophisticated approach (e.g., multiple imputation) [Schaffer, 2005].	1.) Can be used for categorical or continuous predictors. 2.) Is supported by software packages (SPSS has a rather extensive package that deals with missing data.) 3.) Well accepted among statisticians [Rubin, 1976].	1.) Can be challenging for some statistical software packages that are not tooled to address missing data. 2.) Even statistical software packages that claim to apply imputation techniques may not properly adjust variances and may provide inflated estimates of model confidence. 3.) Can produce worse results if applied incorrectly than if not applied at all [Little, 1988].

Table 2.8 provides an overview of some types of optimization approaches that may be appropriate for cases where the cost function is a logistic regression function.

Table 2.8 Optimization Approaches

Approach	Description	Pros	Cons
Iterative optimization approach	This approach involves approximating an objective function using a simpler form (e.g. quadratic) and seeing if the optimal function value at an estimated optima is within a specified tolerance of the value of the true objective function evaluated at that point. If it is, then an approximation for the optima that is good enough has been found. If not, then refine the model form or search in another part of the feasible region and iterate again or modify the approximation of the objective function and iterate again. Both line search and trust region methods [Kelley, 1999] appear to be viable options.	Approaches of this type are often well understood and well documented. Precedents may exist that can be leveraged.	Potential exists for there to be a prohibitive level of computational time involved.
Stochastic multi-stage linear programming	Extension of linear programming where stochastic variables are in the constraint and objective functions.	(1) LP problems are straightforward to formulate and solve. Approach by Dantzig and Infanger [1993] invokes Bender's decomposition (2) Approach appears to be highly flexible	(1) May have a very large state space of variables (2) Objective functions and constraints unlikely to be linear and will require some sort of linearization approximations or piecewise treatment (3) May be computationally intensive
Genetic algorithms	Search algorithms inspired by natural evolution	(1) Have been used for portfolio selection	(1) More efficient and better understood approaches

Table 2.8—*Continued*

		applications [Chan et al., 2002]. (2) Well-suited for finding an approximation for an optimal point on an N-dimensional surface (3) Software exists off the shelf for some genetic algorithm problems	may exist (2) Likely unnecessary if test function not plagued with issues of local extrema
Baseline with adjustments for various different levels of different factors	Analyze what distributions seem to correlate with proposal opportunities where a contract is awarded. Recommend conformity with this profile.	(1) Could be implemented as a vector extension of a model that considers multiple factors. Very straightforward. (2) Precedent exists in software engineering [Yang et al., 2008]	(1) "What winners tends to do" does not necessarily equal "optimal strategy". (2) To rigorously account for confounding of multiple factors will be daunting as the number of factors to be considered grows. (3) Must build-in logic to ensure that results are sensible. In other words, there should be no cases where a negative percentage of effort to be spent on an activity is recommended.

Chapter 3

Research Design

This chapter provides definition of detailed research questions that expand and more precisely define the research question introduced in Chapter 1. Detailed hypotheses, that elaborate on the hypotheses presented in Chapter 1, are also defined. An overview of the research design is presented, and each activity in the research design is discussed. The chapter concludes with study limitations and threats to validity.

3.1 Detailed Research Questions

The primary research question to be addressed by this dissertation is, "When defining the proposed offering to be included in a proposal, what is an optimal use of systems engineering to maximize the probability of being awarded a contract?" In order to begin the optimization process, it is first necessary to identify the relevant variables. In order to begin identifying such variables, characteristics of classes in a systems engineering strategy framework [Smartt and Ferreira, 2012a] were analyzed. Analysis of the framework reveals that the following can be potentially controlled to move an organization from its current state to a desired future state: (1) characteristics of the organization, including the organization's structure, employees, employee skill level, culture and overall process maturity, (2) some characteristics of the environment such as the customer base and potential partners, (3) characteristics of the product or service, and (4) characteristics of the processes employed (including systems engineering processes). Changing the environment can be challenging, and there are limitations as to what degree an organization can change its environment. For example, organizations generally have limited ability to change their regulatory environments [Hutzschenreuter and Israel, 2009]. Different types of products or services are used in proposal management.

Decision makers for a proposal project seek to tailor systems engineering processes such that an optimal level of understanding of customer needs and system definition can be attained within schedule or budget constraints. This research seeks to determine what constitutes an optimal level of systems engineering effort, in terms of the overall budget or number of labor hours devoted to systems engineering, but also in terms of the distribution of labor hours across various systems engineering activities and to employees of varying skill levels.

The focus of this research is at the project level for a particular proposal opportunity. Therefore, more slowly evolving organization characteristics such as organization culture, general structure and system and software process maturity level may be more or less determined at the time a particular proposal opportunity arises. However, people in decision making roles on the proposal project are often able to select which employees will be part of the proposal project team, allocate budget and direct what activities those employees should be doing during the proposal effort.

Figure 3.1 provides a timeline for a proposal effort that is defined to help frame the research questions.

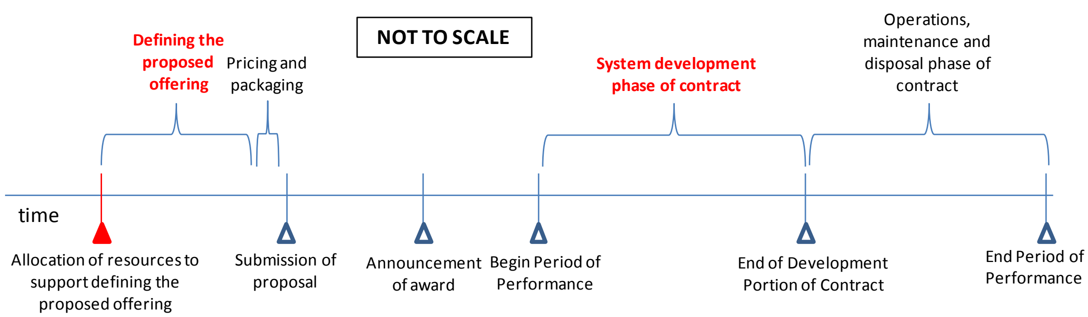


Figure 3.1 Timeline of Interest

Figure 3.1 is simplified to show the development phase of the contract as ending before the operations, maintenance and disposal phase begins. The timeline is not to scale. For example, some systems have an operations and maintenance phase that is many times as long as the development phase. In reality, often systems from the first lots of production are in use while later lots are still being developed. The research questions in this dissertation pertain to decisions that are made at the beginning of the proposal effort (red solid triangle) related to resource allocation during the portion of the proposal effort that involves defining the proposed offering. These allocation decisions depend upon a number of factors, including an initial understanding of how much development effort exists in the system development phase of the contract. This estimate of contract size used to make proposal effort resource allocation decisions is likely crude as it must be determined before any detailed exploration of the system definition. A standard metric must be defined to quantify project size, so that projects of different sizes can be compared. Potential options for such a metric include total revenue from developing the system, number of units of the system or number of labor hours spent working on the system. After considering many options, the size metric selected to normalize data in this research is the number of engineering labor hours to be spent on the development portion of the contract. This metric was chosen somewhat arbitrarily, but it is conducive to understanding the systems engineering effort on the proposal as a fraction of the engineering effort on the development portion of the contract.

3.1.1 Systems Engineering Factors and Proposal Success

One important question that this research seeks to answer is "What systems engineering related factors impact the probability of contract awards?"

After analyzing the attributes of the classes in the systems engineering strategy framework and synthesizing feedback from various industry experts, the following set of factors are explored as factors that may potentially impact the probability of a contract award. These include: (1) level of system and software process maturity, (2) motivation for submitting the proposal, (3) number of labor hours spent on the proposal performing systems engineering activities related to defining the proposed offering normalized by contract size, (4) extent of ongoing or past relationships with a particular customer, (5) where past or ongoing relationships exist, perceived level of customer satisfaction with previous or ongoing contract efforts, (6) type of contractual arrangement, (7) number of interactions with the customer normalized by contract size, (8) competitive rank of the organization and (9) experience level of proposal project leadership. For a more precise definition of these factors and various levels of each factor see Table D.4.

3.1.2 Optimal Total Level of Systems Engineering Effort

This section provides a detailed breakdown of the question related to the optimal total level of systems engineering effort defining the proposed offering. This research seeks to answer the following: "When defining the proposed offering to be included in a proposal, how does the ratio of *total labor hours spent performing systems engineering activities defining the proposed offering on the proposal effort* to *the estimated total number of engineering labor hours to be worked on development portion of the contract* impact the probability of being awarded a contract?"

3.1.3 Relative Allocation of Labor Hours for Various Systems Engineering Activities

This section addresses the issue of allocating the total labor hours for systems engineering between various activities. The activities considered in this research are those defined in Table 2.1: identify source requirements, validate source requirements, identify system requirements, validate system requirements, define architecture alternatives, validate architecture alternatives (includes identifying and evaluating key technologies), estimate cost for each alternative and perform decision analysis to select a preferred solution.

It is expected that decision makers may in some cases re-allocate effort to various activities numerous times throughout the proposal effort depending upon the particular challenges and opportunities that arise. Nonetheless, it is helpful for planning (i.e., what skill sets need to be on a team) to have an estimate of how the labor hours will be divided between the activities and skill levels. The research question is: "When defining the proposed offering to be included in a proposal, what is the optimal percentage of the systems engineering labor hours to expend on each of the various SE activities to maximize the probability of being awarded a contract?"

3.1.4 Relative Allocation of Labor Hours Across Various Skill Levels for Each Activity

For each proposal activity, it is useful to determine the optimal mix of skill levels. Optimality in this sense considers both the productivity achieved by the contributors and the hourly labor costs of the time these contributors invest. For example, if a very senior person is

20% more productive than a much less senior person but costs 30% more, it may be optimal to allocate the majority of the budget to the less senior person and use the more senior person only in situations where experience is required. The primary research question related to skill level distribution is:

"For each of the various SE activities that are part of defining the proposed offering, what is the optimal distribution of budget and labor hours to individuals with various skill levels to maximize the probability of being awarded a contract within prescribed budget constraints?"

In order to effectively answer this question it is important to define a scale to describe skill level. A number of employee attributes must be considered when defining skill level. As systems thinking is a key skill for systems engineers of all types [INCOSE, 2010], it is reasonable to assume that attributes correlated with high systems thinking skill levels may in fact also be correlated with high systems engineering skill levels. Attributes identified as potential enablers for systems thinking skills include: work experiences, education, individual characteristics, life experiences outside of work, interpersonal skills, and training [Davidz and Nightingale, 2008]. Another perspective about what qualities impact skill levels of systems engineers can be found in INCOSE's criteria for certifying systems engineers. INCOSE's certification program recognizes experience, education, performance on a standardized exam and input from references [INCOSE website, 2013]. Because systems thinking skills, individual characteristics, life experiences outside of work, interpersonal skills, specific training history, likely score on a standardized exam that is not available for organizations to administer and input from references are all difficult to objectively assess without an extensive investigation, the skill rubric defined in this dissertation will focus primarily on work experience and education. Two categories of work experience will be considered: work experience in the particular activity and experience working in the problem domain. It is expected that different problem domains have different levels of technical depth and that a number of years that may render an employee a domain expert in one domain may not qualify that same employee as an apprentice in another domain. Table 3.1 provides a definition of the four levels used throughout this dissertation. The four levels are selected for a reason. To some degree, systems engineering is a craft. Therefore, when selecting levels, it seemed that an equivalent was needed for apprentice, journeyman and master. The equivalent levels defined here are labeled “beginner”, “intermediate” and “advanced” respectively as to not overly belabor the craft analogy. However, there is a need for an additional skill level beyond the master level for truly expert knowledge. Therefore, an additional level of “expert” is added.

Table 3.1 Skill Levels Scale

Skill Level	Formal Education Level	Number of Years Domain Experience	Number of Years of Experience Performing Particular Activity
Expert	No college degree	20+ years	6+ years
	Associates degree or bachelor’s degree in non-technical field	16+ years	6+ years
	Bachelor’s degree in technical field	12+ years	6+ years
	Masters or PhD in technical field	8+ years	6+ years
Advanced	No college degree	16+ years	6+ years
	Associates degree or bachelor’s degree in non-technical field	12+ years	6+ years
	Bachelor’s degree in technical field	8+ years	6+ years
	Masters or PhD in technical field	4+ years	4+ years
Intermediate	No college degree	12+ years	6+ years
	Associates degree or bachelor’s degree in non-technical field	8+ years	6+ years
	Bachelor’s degree in technical field	4+ years	4+ years
	Masters or PhD in a technical field	2+ years	2+ years
Beginner	Does not yet meet requirements for intermediate level		

Each row within a block associated with a skill level provides a set of criteria that all must be met to qualify an individual at that level. If an individual meets the set of criteria on any row within a block, that individual can be classified accordingly. For example, someone with a bachelor’s degree in a technical field, 9 years of experience in a problem domain, all of which is devoted to architecture, is at an advanced level for architecture. However, that same individual is at a beginner level for all other systems engineering activities considered.

The particular number of years of experience for each skill level is somewhat arbitrarily defined as is the relative value of formal education vs. experience. In some problem domains, the level of formal education may be critically important and number of years of experience less so. In other problem domains, an advanced level of education may not be required, but years of experience are critical. Nonetheless, some defined criteria are necessary to support the collection of consistent results. As each of the systems engineering activities identified in Table 2.1 is different, it is reasonable to assume that there exists a unique skill level breakout for each activity. Therefore, the following questions must be answered:

A.) "When defining and validating source requirements in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating source requirements?"

B.) "When defining and validating system technical requirements in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating system technical requirements?"

C.) "When defining and validating system architecture concepts in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating system architecture concepts?"

D.) "When estimating costs of various candidate system solution concepts in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in cost estimation?"

E.) "When performing decision analysis to select a final solution in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in performing decision analysis?”

It is expected that the answers to these questions will be highly dependent on the particular organization and the product and processes involved. While a survey is conducted to collect data in an attempt to address these questions, an approach is desired that provides a way to answer the questions for specific proposal opportunities in particular organizations. To provide this, a framework is discussed for developing a decision support system to help decision makers optimize the use of systems engineering on proposals within their organization.

3.1.5 Summary of Research Questions

Each of the detailed research questions is provided in Table 3.2. This is provided as a quick reference to help readers extract the actual questions from the background and explanatory information presented in the previous sections.

3.2 Detailed Hypotheses

Two hypotheses were presented in Chapter 1:

(1) Investing in a higher number of systems engineering labor hours to support the proposal process will result in a higher probability of contract award than investing in a lower number of systems engineering labor hours.

(2) Executing proposal efforts with particular factors at specific levels (i.e., more sophisticated systems engineering process maturity level, greater level of customer satisfaction on past contracts, a higher number of exchanges with the customer prior to the proposal submission, elevated competitive rank relative to the competition) will positively impact the probability of contract award.

Table 3.2 Summary of Detailed Research Questions

Question
What systems engineering related factors impact the probability of contract awards?
When defining the proposed offering to be included in a proposal, how does the ratio of *total labor hours spent performing systems engineering activities defining the proposed offering on the proposal effort* to *the estimated total number of engineering labor hours to be worked on development portion of the contract* impact the probability of being awarded a contract?
When defining the proposed offering to be included in a proposal, what is the optimal percentage of the systems engineering labor hours to expend on each of the various SE activities to maximize the probability of being awarded a contract?
When defining and validating source requirements in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating source requirements?
When defining and validating system technical requirements in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating system technical requirements?
When defining and validating system architecture concepts in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating system architecture concepts?
When estimating costs of various candidate system solution concepts in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in cost estimation?
When performing decision analysis to select a final solution in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in performing decision analysis?

The additional definition and introduction of background concepts allows these hypotheses to be more precisely defined. In the case of the first hypothesis, the hypothesis is redefined in more precise terms and labeled 1-A to differentiate it from the more broadly stated Hypothesis 1. In the case of the second hypothesis, the hypothesis is decomposed into a number of singular hypotheses, each of which can be evaluated. These more detailed hypotheses are described in Table 3.3. A rationale is also provided for why the author believes each of these hypotheses to be potentially true. As no empirical research was found that examined these factors in the context of the probability of contract award prior to this dissertation research, the author's experience as well as general principles taught in various technical and managerial courses attended by the author form most of the rationale. Observe that there is not a one-to-one mapping between detailed research questions and detailed hypotheses. For some detailed research questions, no detailed hypotheses were formulated because there was no a priori notion of the relationships between the variables. For other detailed research questions, multiple hypotheses were formulated because there were multiple sets of relationships between variables for which a priori notions of the relationships existed.

3.3 Overview of Research Design

The research design is the sequence of research activities that are undertaken to evaluate the research hypotheses and address research questions. Figure 3.2 provides an overview of the research design planned for this dissertation.

The research design begins with defining use cases for a DSS designed to optimize the use of systems engineering on a particular proposal project. A high level architecture is defined, data needs are identified, a survey is created and administered to gather data, and a factor analysis is performed on the survey results.

Table 3.3 Detailed Hypothesis Statements and Rationale

Hypothesis Number	Hypothesis Statement	Rationale for Hypothesis
1-A	As the ratio of number of systems engineering labor hours defining the proposed offering to the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	Spending more time understanding source and system requirements, exploring architectures, examining possible technology insertion options, carefully estimating costs and methodically comparing various candidate solutions to select the best value solution should result in a better proposed offering.
2-A	As the number of written interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	More contact with a customer provides more opportunity for the organization to understand the customer's needs and convince the customer that the organization can fulfill those needs.
2-B	As the number of telephone interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	More contact with a customer provides more opportunity for the organization to understand the customer's needs and convince the customer that the organization can fulfill those needs.
2-C	As the number of video conference interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	More contact with a customer provides more opportunity for the organization to understand the customer's needs and convince the customer that the organization can fulfill those needs.
2-D	As the number of in-person interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	More contact with a customer provides more opportunity for the organization to understand the customer's needs and convince the customer that the organization can fulfill those needs.
2-E	As the level of systems and software process maturity of the prime contractor increases, the probability of contract award increases.	An organization holding a high level of system and software process maturity is attractive to customers. A high process maturity rating is a general indicator of mature processes. Research has linked mature systems engineering processes to potential programmatic benefits [Elm, 2007].
2-F	As the perceived level of customer satisfaction with previous	Customers generally prefer to do business with

Table 3.3 —*Continued*

	or ongoing contract efforts increases, the probability of contract award increases.	organizations that have done highly satisfactory work for them in the past.
2-G	As the competitive rank of the prime contractor with respect to market share increases, the probability of contract award increases.	Name recognition sells.
2-H	As the competitive rank of the prime contractor with respect to prestige increases, the probability of contract award increases.	Prestige sells.
2-I	As the experience level of the project manager of the proposal project for the prime contractor increases, the probability of contract award increases.	An experienced project manager on a proposal will be more adept at managing the interface with the customer throughout the proposal process and ensuring that the customer's needs are being effectively addressed than a less experienced project manager.

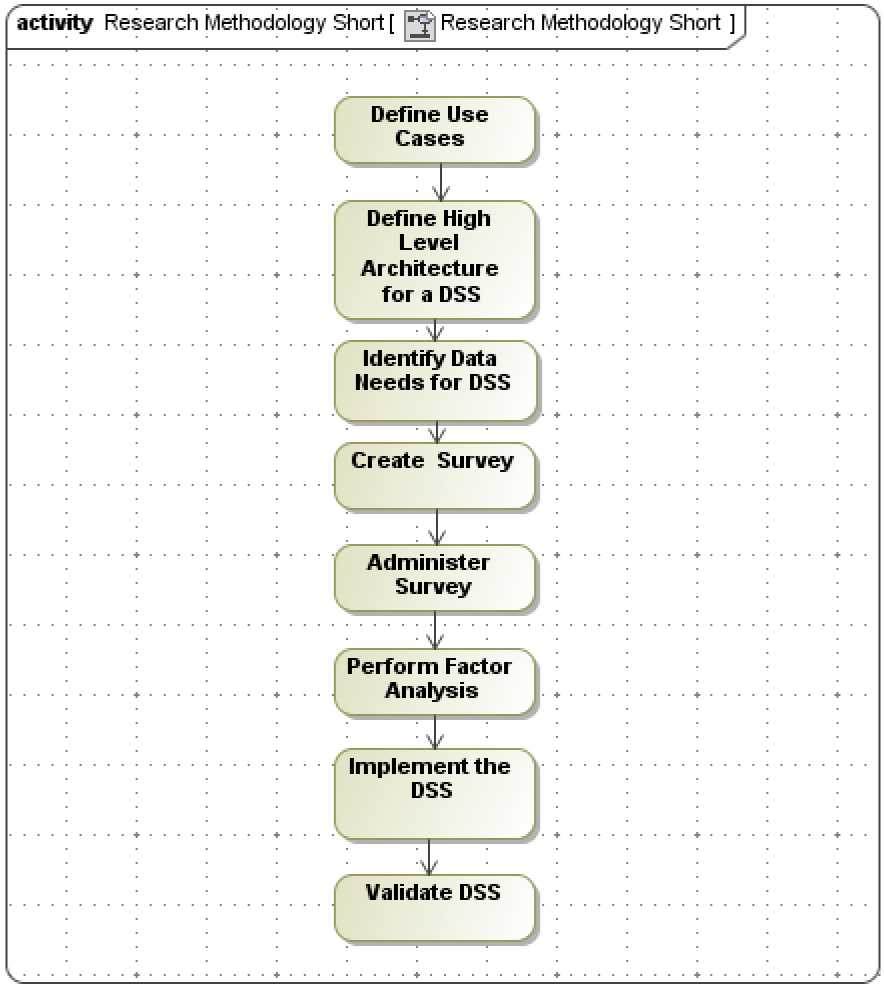


Figure 3.2 Research Design Overview – As Planned

The number of survey responses was inadequate to proceed with developing a DSS with parameter values derived from the survey data (see Section 4.1.2.3 for details). Therefore, the research design had to be modified so that instead of there being a DSS, an optimization modeling framework is introduced instead. It is believed that some organizations may have an adequate quantity of historical proposal efforts to be able to derive parameters for a DSS, and this framework can be used in the future to help such organizations develop decision support systems. With such future decision support systems, decision makers can find optimal resource allocation solutions to their particular proposal efforts that consider characteristics unique to the proposal effort.

Figure 3.3 displays an overview of the research design as implemented. The first several steps in the research design that was implemented are identical to the first several steps in the research design that was planned. However, “Implement the DSS” has been replaced with “Develop Optimization Modeling Framework” and “Validate DSS” has been replaced with “Validate Optimization Modeling Framework”. The next several sections of this chapter are organized in relationship to the research design that was implemented.

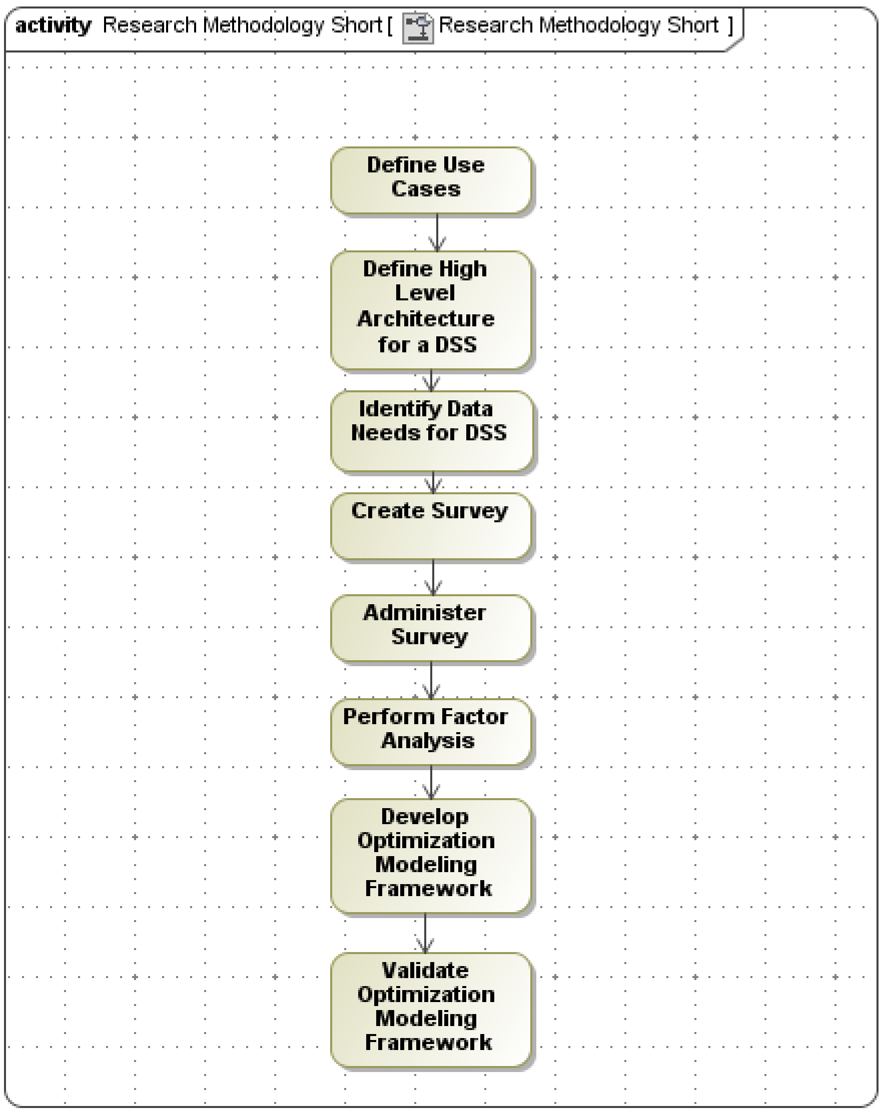


Figure 3.3 Research Design Overview – As Implemented

3.4 Defining a Use Case for the DSS

In order to design and develop a DSS, it is first necessary to determine who the DSS is supposed to serve and what the DSS is supposed to do for those it serves. To accomplish this, a use case analysis of the DSS is performed.

3.4.1 Use Case Diagram

Figure 3.4 provides a use case diagram that describes the users and the use case that the DSS will support. Each of these users and the use case are explored in detail in a later section.

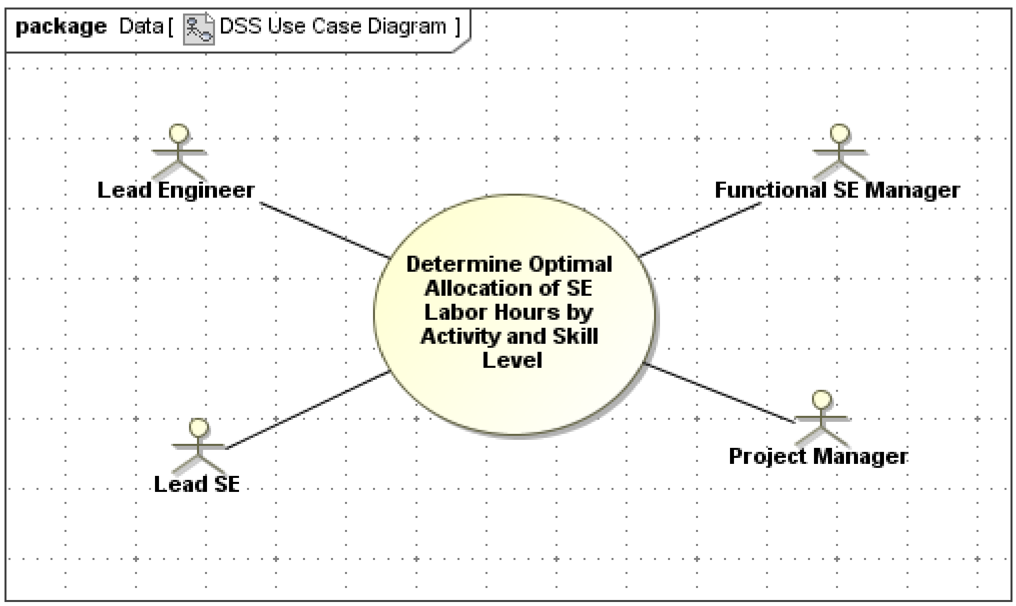


Figure 3.4 DSS Use Case Diagram

3.4.2 Users of the DSS

This section describes the set of users who will use the DSS. Table 3.4 provides an overview of the intended users of the DSS. Note that the set of users is a subset of the stakeholders for the proposal process identified in Table 2.3. Company shareholders, top executives, program managers, capture managers and individual contributors are not included in the set of intended users of the DSS. While company shareholders and top executives are clearly beneficiaries of successful proposal management, in most organizations they would not be involved in making the types of systems engineering resource allocation decisions that the DSS facilitates. Program managers would likely create a project for the proposal effort and delegate responsibility for the allocation of resources on the proposal to an assigned project manager. Capture managers would likely be more focused on the interface with the customer and less focused on the processes to define the proposed offering. Individual contributors also generally do not allocate resources or assign personnel to work on proposal efforts.

Table 3.4 Users of the Decision Support System

User	Definition
Project Manager	"The person appointed to take day-to-day responsibility for management of the project throughout its stages" [Leach, 2000].
Lead Engineer	The lead engineer coordinates the engineering aspects of a project. This individual interfaces with both project management and the functional managers of the organizations responsible for systems engineering expertise.
Lead Systems Engineer	The lead systems engineer coordinates the systems engineering effort on a project. This individual interfaces with project management, the lead engineer, and functional managers of the organization providing systems engineering expertise. The lead systems engineer is also responsible for ensuring that the technical effort is conducted in an integrated way. This includes establishing and managing the integrated product and process teams. In some cases, there exists one lead engineer for an organization that may engineer an entire portfolio of systems. In such cases, the lead systems engineer may be the highest ranked engineer dedicated to the project and may effectively act as the lead engineer for the project on a day-to-day basis.
Functional Manager of Organization Performing Systems Engineering	In matrix organizations, there are often organizations chartered to perform a particular function. Employees are often managed by these functional organizations and assigned to work particular projects. The managers of these functional organizations generally assign particular individuals to work on particular proposal efforts. The functional management also generally defines the processes by which work is to be executed.

3.4.3 Use Case Definition

This section will provide additional definition for the use case Determine Optimal Allocation of SE Labor Hours by Activity and Skill Level. Table 3.5 provides a detailed definition of this use case.

Table 3.5 Definition: Determine Optimal Allocation of SE Labor Hours by Activity and Skill Level

Use Case Name:	Determine Optimal Allocation of SE Labor Hours by Activity and Skill Level
Goal:	Determine the optimal number of systems engineering labor hours to allocate to each activity (i.e., source requirements, system technical requirements, system architecture, cost estimation, and decision analysis) and each skill level (i.e., beginner, intermediate, advanced, expert) to maximize the probability of a contract award within a given budget
Summary:	Determine optimal plan for how to use systems engineering to define the proposed offering for a proposal
Actors:	1. Project Manager 2. Lead Engineer 3. Lead Systems Engineer 4. Functional Manager of Organization Responsible for Systems Engineering
Trigger:	Actor selects to initiate use case.
Preconditions:	1. The organization has at least initially decided to pursue a particular contract. 2. Fulfilling the contract will require the development of at least one complex system. 3. There is a need for systems engineering to be used to understand the customer's requirements and to define a solution. 4. A statistical characterization of the factors has been performed such that an equation has been created that expresses probability of contract award as a function of how much systems engineering labor is allocated to particular activities and employees with particular skill levels as well as potentially other relevant factors. 5. The following proposal information is available: a. Contract (estimated total number of engineering hours to be performed on the contract, industry contract is being performed in, type of contractual arrangement) b. SE organization (system and software process maturity level, the maximum number of available labor hours for contributors at various skill levels to perform each of the systems engineering activities in the proposal effort, hourly labor rates for contributors at different skill levels for each systems engineering activity, level of project manager experience, motivation for submitting the proposal) c. Competition (competitive rank with respect to market share, competitive rank with respect to prestige) d. Customer familiarity (extent of past or ongoing relationships, level of customer satisfaction on previous or ongoing project efforts) e. Proposal effort (number of contacts of various types between customer and prime contractor during proposal effort)
Post conditions:	1. Actor knows how many hours to allocate per activity and skill level to define the proposed offering. 2. Actor knows the expected probability of contract award (*P*) and expected cost associated with the recommended allocation.

Table 3.5 —*Continued*

Normal Flow:	1. Actor inputs proposal information and budget not to be exceeded. 2. DSS calculates a maximum *P* profile based on proposal inputs. 3. DSS displays *P* profile to the actor. 4. Actor selects level of investment that presents the best probability of contract award per cost value. 5. DSS displays the optimal number of labor hours for each activity and skill level.
Alternative Flow:	1. Actor inputs proposal information. 2. DSS calculates a maximum *P* profile based on proposal inputs. 3. DSS displays *P* profile to the actor. 4. For a subset of values from *P* profile, DSS displays the optimal number of labor hours for each activity and skill level.

3.4.3.1 Example Scenario – Input/Output Dialogue Between DSS and User

The following discussion provides an alternative view of the user's interaction with the DSS. This viewpoint represents an idealized vision of the DSS in which sufficient historical data is available to allow for the DSS to provide detailed guidance on how to allocate resources considering both activity and skill level. If less historical data is available, alternative formulations can be developed to provide as much detailed guidance as is possible with the quantity of available historical data. In the normal flow, a user will interact with the DSS via a sequence of inputs and outputs. A decision maker must decide how much budget to allocate to systems engineering labor to support defining the proposed offering based upon the cost of the labor and an estimated probability of contract award.

The following dialogue describes a series of steps that are undertaken when using the DSS. These diagrams are illustrative examples and not based on actual proposal data. These are included to provide a view of what the decision support system inputs and outputs may look like. The reader should not attempt to apply actual values from these example model outputs to actual proposal efforts.

1. The user inputs information about the contract, the SE organization, the competition, the customer, the proposal effort and the budget not to be exceeded (see use case definition in Table 3.5 for details).
2. The DSS runs an optimization for a number of different levels of investment in the proposal effort from no investment up to the maximum budget level specified.

3. Based on the optimal allocation strategy (by skill level and activity) given each constraint, an estimated probability of contract award vs. level of investment in systems engineering labor is presented (Figure 3.5):

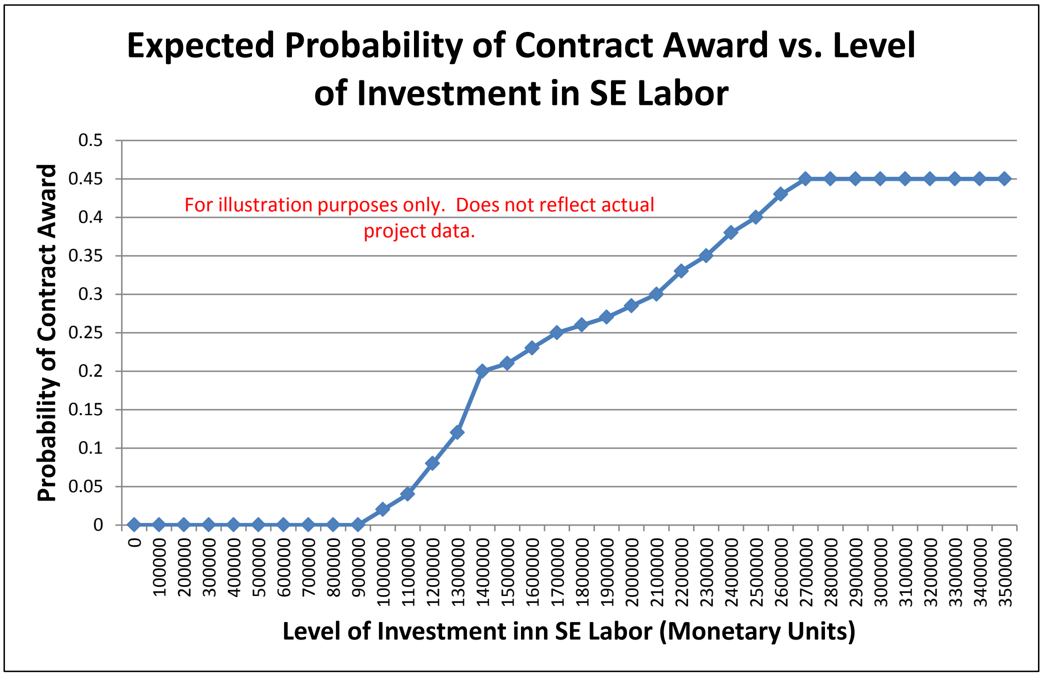


Figure 3.5 *P* vs. Level of Investment

4. The user selects a budget.

5. The DSS then presents the optimal allocation of systems engineering labor given by skill level and activity. An example of this output is displayed in Figure 3.6.

The DSS will allow the user to perform "what if" analysis. For example, the DSS can be run multiple times with different budgets to see the potential benefits of increasing the budget. Also, the DSS can be run multiple times to allow an organization to explore potential impacts of changing levels for variables that are static in the model. For instance, an organization may wish to use the DSS to understand how investing in increasing the system and software process maturity levels of the organization might impact the probability of a contract award.

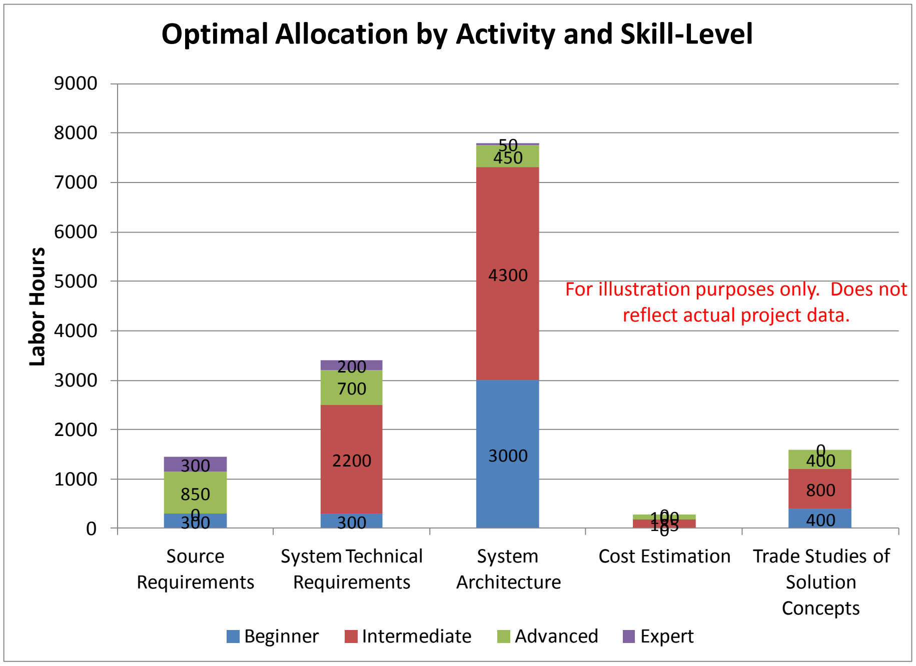


Figure 3.6 Example Recommended Allocation Scheme

3.5 Defining the High Level Architecture for the DSS

This section defines the architecture for the DSS. The ultimate goal of the architecture analysis is to define the architecture for the DSS in sufficient detail to guide the implementation. This section presents a context diagram and a functional block diagram to provide high-level views of what functionality is in the DSS.

3.5.1 DSS Context Diagram

Figure 3.7 displays a context diagram for the DSS.

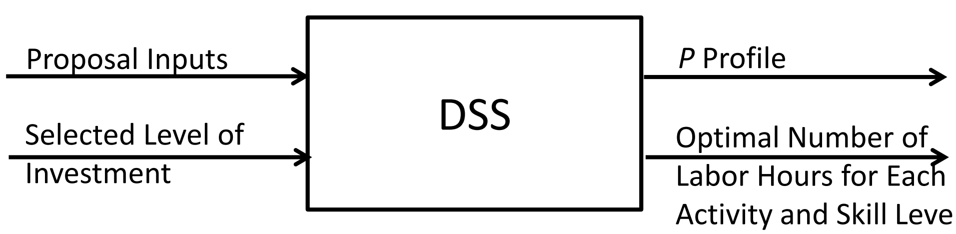


Figure 3.7 DSS Context Diagram

The inputs to the DSS are the proposal inputs and a selected level of investment. The proposal inputs include information about the contract, SE organization, competition, customer and proposal effort. The selected level of investment is the amount of budget that the user determines provides the best tradeoff between level of investment in systems engineering on the proposal and the probability of contract award.

The outputs of the DSS are a *P* Profile and a recommendation for the optimal number of labor hours for each activity and skill level. The *P* Profile is a set of probability of award estimates as a function of level of investment in systems engineering labor to define the proposed offering. Figure 3.5 provides a graphical depiction of the *P* Profile. The recommendation for the optimal number of labor hours for each activity and skill level is graphically depicted in Figure 3.6. As described in the alternative flow in the use case (defined in Table 3.5), it is possible the DSS provides an optimal number of labor hours distribution for every data point in the *P* Profile. When this is the case, it is unnecessary for a user to input a selected level of investment.

3.5.2 DSS Functional Architecture Block Diagram

Figure 3.8 provides a functional architecture block diagram for the DSS. The functional architecture block diagram provides greater insight into what functionality is internal to the DSS.

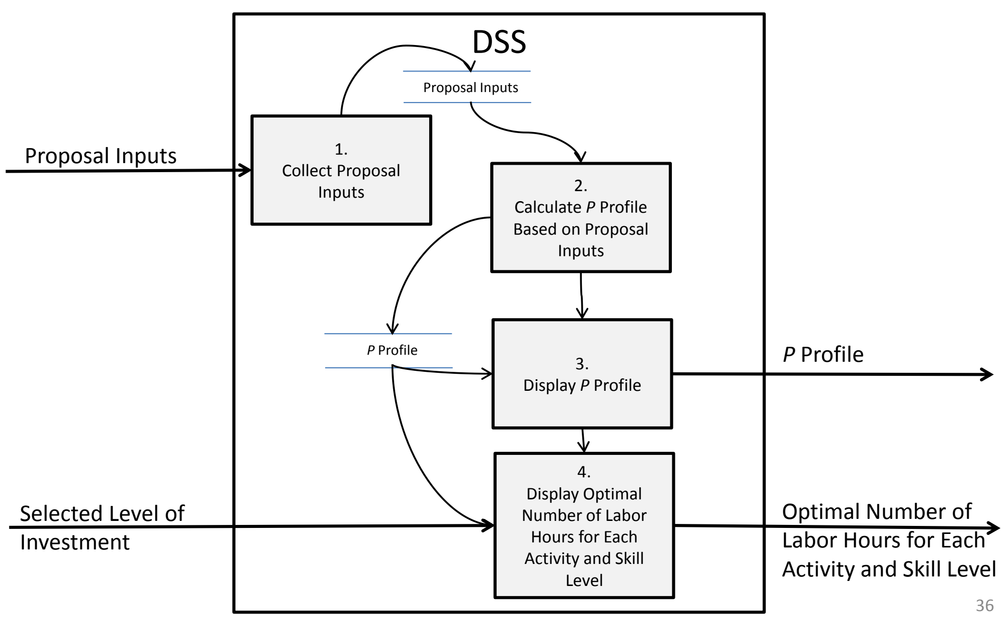


Figure 3.8 DSS Functional Architecture Block Diagram

Note that the inputs and outputs are the same for the functional block diagram as for the context diagram. Additional definition is provided for the major functions within the DSS. The major functions are: Collect Proposal Inputs, Calculate *P* Profile Based on Proposal Inputs, Display *P* Profile, and Display Optimal Number of Labor Hours for Each Activity and Skill Level. Collect Proposal Inputs is a function of the DSS that enables the user to input information about the proposal effort into the DSS.

Calculate *P* Profile Based on Proposal Inputs calculates the probability of contract award for a subset of the possible levels of investment. For each level of investment, the value of *P* reflects an optimal allocation of systems engineering labor hours for activities and skill levels. Therefore, as the *P* Profile is being calculated, the optimal labor allocation for each level of investment in the *P* Profile is also being calculated. The *P* Profile and the related labor allocation schemes are stored while the DSS is running. Calculate *P* Profile solves an optimization problem for each level of investment. This optimization model is formalized in mathematical programming language in the Section 3.5.3.

Display *P* Profile and Display Optimal Number of Labor Hours for Each Activity and Skill Level are two functions that allow the user to view the information that was calculated in Calculate *P* Profile Based on Proposal Inputs. Display *P* Profile actually displays the *P* Profile to the user and provides the user with information that allows the user to select the level of investment for which the resource allocation recommendation will be viewed. Display Optimal Number of Labor Hours for Each Activity and Skill Level actually provides the user with information related to the optimal number of labor hours to allocate to each activity and skill level when defining the proposed offering. Once again, an alternative formulation is possible where the DSS provides an optimal number of labor hours distribution for every data point in the *P* Profile. When this is the case, it is unnecessary for a user to input a selected level of investment.

3.5.3 Mathematical Description of Optimization

The optimization problem in Calculate *P* Profile Based on Proposal Inputs involves maximizing an objective function *P* over a feasible region defined by a set of constraints. The set of constraints is dependent on the particular proposal effort of interest. In other words, the DSS is designed to make recommendations that accommodate the particular set of constraints of a particular proposal project. For example, if there are no expert-level contributors for source requirements, this can be represented as a constraint and the DSS will recommend a labor allocation that does not call for the use of expert-level contributors.

The optimization function of the DSS can be described in mathematical programming language. The objective function and three constraints are listed below.

Objective Function:

$$z(B) = \max(\hat{P}(x))$$

Constraints:

$$x_{ij} \le u_{ij} \quad \text{for all } i \text{ and } j \tag{1}$$

$$x_{ij} \ge 0 \quad \text{for all } i \text{ and } j \tag{2}$$

$$\sum_i \sum_j t_{ij} x_{ij} \le B \tag{3}$$

where:

x_{ij} is the number of systems engineering labor hours for contributors at skill level *j* on activity *i*

u_{ij} is the maximum number of systems engineering labor hours that are available for skill level *j* on activity *i*

t_{ij} is the hourly rate for contributors at skill level *j* on activity *i*

B is the budget or total monetary value not to be exceeded

$\hat{P}$ is the estimated probability of contract award

The function P (the true probability of contract award) is expressed as a binary multivariate logistic regression function of a number of variables. A logistic regression function is used because the output values are between 0 and 1. Logistic regression functions never assume a value of 0 or 1, but approach these values asymptotically. This property of logistic regression functions is well-aligned with the reality of making investments in systems engineering labor on proposals. Even if the available budget for systems engineering labor approaches infinity, there is generally no guarantee of a contract award. Key independent variables relate to the ratio of hours spent on the proposal effort to the estimated number of labor hours on contract if the contract is awarded. These variables are defined below.

$$\frac{x_{ij}}{D}$$

where:

x_{ij} is the number of systems engineering labor hours for contributors at skill level j on activity i

D is the projected number of total engineering hours on contract for the development portion of the system lifecycle if awarded

There may be other candidate factors that impact P but are not under the control of the decision maker when optimizing the use of systems engineering on proposals. The set of such factors considered in this research is: (i) level of system and software process maturity of respondent organization, (ii) level of system and software process maturity of the prime contractor if the organization is a subcontractor on a team competing for a contract, (iii) level of system and software process maturity on the proposal effort project, (iv) level of project manager experience, (v) competitive rank of the prime, (vi) extent of the relationship with the proposal customer, and (vii) level of customer satisfaction on previous or ongoing contract efforts. Observe that u and t are not factors in P, but rather proposal-effort specific constraints that help define the feasible region. Because the values for u and t constrain the options, they can impact the optimal solution even if not represented directly in the objective function.

3.5.4 Validating the Use Cases and Architecture

Prior to conducting the survey and developing example decision support systems, the use cases, input/output sequence, elements of the architecture, mathematical programming formulation of the optimization problem at the heart of the DSS and scheme for implementation were reviewed multiple times by dissertation committee members and refined. As a result, this information was also provided to UTA's Systems Engineering Research Center (SERC)'s Industry Advisory Board (IAB) for feedback. The UTA SERC IAB is composed of leaders from area organizations that engineer complex systems. Many of the IAB members are highly knowledgeable about the use of systems engineering on proposals.

After the survey was conducted and example decision support systems were developed, further validation of the architecture was performed. Industry experts from several organizations who engineer complex systems each were presented with a presentation of the DSS framework designed to take one hour including questions and answers. Part of the presentation related to the architecture of the decision support system and the formal mathematical programming definition of the optimization model at the heart of the decision support system.

3.6 Identifying Data Needs for the DSS

The survey is designed to collect data related to the use of systems engineering in proposal management, data to understand the demographics of the respondents and proposal efforts, and data to validate survey responses. The required data items are: (1) data items necessary to answer one or more research question, (2) demographic data items designed to assess other potentially important characteristics of the organization, environment, product and processes related to a particular proposal, (3) data items relating to the award status of the proposal opportunity, and (4) data to validate the survey responses.

The following activities were performed to determine a set of data items. First, the research questions were carefully analyzed to determine what data needs to be collected to address them. Second, all attributes of the various classes of information in the systems engineering strategy framework [Smartt and Ferreira, 2012b] were analyzed to determine which attributes were applicable to the use of systems engineering in proposal management. Third, other models from industry designed to make decisions related to proposals were examined to identify additional pertinent data items. Finally, the consolidated set of the data items was presented to industry experts. A few additional data items were identified as a result of this expert input.

3.7 Creating the Survey

This section relates to the process of creating the survey.

3.7.1 Selecting Survey Items

Survey data items are designed to (1) collect information related to the data needs of the survey, (2) gain important demographic information on the respondents, contracts being pursued and the proposal efforts, and (3) evaluate whether a survey response is valid and within the scope of analysis.

3.7.2 Formulating the Survey Questions

Each of the survey items is measured through one or more specific survey questions. A number of demographic questions are also included on the survey. Each question is designed to elicit exactly one data item. The principles outlined for writing effective survey questions discussed in Section 2.2.1.3 were used when constructing survey questions.

Many sources on survey methods [Fowler, 1993; Sheatsley, 1983] recommend presenting the survey questions and survey instrument to multiple knowledgeable individuals and asking them to attempt to take the survey prior to using the survey to collect data. Through this process, the individuals usually find items in the survey that are unclear and ask questions. This also provides additional opportunity for any errors, such as grammatical errors, to be corrected before the survey instrument is used. Inputs from other people help the researcher identify survey items that are not clear. In addition, the researcher may find that these respondents interpret questions in ways that are accurate with respect to how the question was asked, but not in the way that was intended. In such cases, the researcher then has an opportunity to craft the question to provide clearer directions about how the respondent should address the survey question.

For this survey, the survey questions were reviewed by all members of the dissertation committee, two additional subject matter experts, four peer reviewers and five industry experts asked to field test the survey. The reviewers were asked specifically to identify any questions that were unclear and to flag questions for which the reviewer could identify more than one way the question could be interpreted. As a result of the feedback, a number of questions were added or modified. A more detailed synopsis of the survey review activity is captured in Appendix C.

3.7.3 Creating the Survey Instrument

Two versions of the survey were implemented. After reviewing objective comparisons of survey tools for creating online surveys [Idealware, 2012] and studying the implementation of a relevant survey used in software engineering research [Ferreira, 2002], the decision was made to implement the survey as a custom developed html form to be populated on the internet. Microsoft Expressions was used to generate the basic survey form, and some patching was done to provide specific functionality (e.g., pop-up box that verifies users wish to clear the survey form if they hit the button to clear the form). Specialized software was used capture respondent inputs upon submission and record their information to a database and create and send an email containing their information to a designated account. A commercial web hosting site was used to host the survey as such an option provided the authors with required software. Both the email account receiving the responses and the database were password protected. However, the first deployment of the survey did not capture an adequate number of responses for analysis.

An analysis of feedback from various contacts revealed that the internet firewalls in organizations that hire the types of employees targeted for the survey prohibited access to the Microsoft Expressions version of the survey. It is believed that many large corporations are now migrating to a white-list philosophy of internet management where access to various internet domains is granted one site at a time. It is believed that the domain purchased to host the survey was not on this access list in many organizations. An informal dialogue and testing conducted by contacts who work in such organizations revealed that Survey Monkey, an off-the-shelf tool for creating and hosting surveys, was widely accessible. As a result, the survey was also implemented in Survey Monkey.

3.7.4 Verifying and Validating the Survey

A number of activities were conducted to verify and validate the survey. These included the construction and execution of a detailed verification test plan for both the Microsoft Expressions and the Survey Monkey version of the survey. Peer testers, dissertation committee members, field testers and expert reviewers were also involved. The input of these individuals proved very helpful in refining the survey. Appendix B provides more detail related to verifying the survey, and Appendix C provides more detail related to validating the survey.

3.8 Administering the Survey

Two professional societies helped promote the survey: the International Council on Systems Engineering (INCOSE) and the National Defense Industrial Association (NDIA). INCOSE's mission is to "share, promote and advance the best of systems engineering from across the globe for the benefit of humanity and the planet" [INCOSE website, 2013]. NDIA is the leading professional association focused on the United States defense industry and aims to promote a rigorous, responsive national security team composed of government and contractors [NDIA, 2013]. Research to make the proposal process more effective is of interest to many NDIA members.

Both INCOSE and NDIA supported the survey by providing links to the survey on their websites. INCOSE provided a link on the home page, and NDIA provided a link on a page devoted to systems engineering studies and publications. NDIA also sent an email to members of their Systems Engineering Division. INCOSE's North Texas Chapter sent an email to its membership announcing the survey and requesting membership participation. The researcher also sent emails to the subset of INCOSE members whose contact information was available on

INCOSE's online member directory [INCOSE website, 2013] requesting participation. The researcher also posted a promotion on INCOSE's Linked In page and procured an advertisement in INCOSE's November 2012 ENote. More details about how the survey was promoted are documented in Appendix D.

3.9 Performing Factor Analysis

Once the survey results were collected, then the results were analyzed. Before statistical analysis of the factor relationships began, the individual responses were analyzed for completeness and consistency. Only responses where the respondent provided name and contact information were considered. A few additional consistency checks were performed to validate the survey responses. Appendix D describes a detailed rubric that was used to determine whether each survey response was considered valid for further data analysis. In addition, the overall set of survey results was scanned to see if it appeared that multiple respondents were reporting the same proposal effort.

For responses that were not disregarded, but where some the questions were unanswered, the questions that were answered were considered. In some cases, important parameters for analysis purposes may be calculated by functions of responses to multiple questions. For these items, the respondent had to answer all of the necessary questions that yield data that is used in these calculations in order for the response to be considered for that particular analysis item.

The survey targets an audience that encompasses a wide range of types of proposal efforts, and therefore, there are outliers in the data. Some of these outliers are possibly caused by a respondent not understanding a question and providing incorrect information, but others may reflect legitimate data values that are well outside the central distribution. In the former case, the data needs to be eliminated from the data set. In the latter case, understanding the outlying data point may yield significant insight into the key relationships related to the use of systems engineering in proposal management. Techniques such as box plot analysis, computing the leverage values, analyzing residuals and computing Cook's Distance were used to identify outliers.

Unfortunately, these techniques cannot definitively ascertain whether an outlier is a legitimate member of the data set or not. Therefore, the researcher used discretion as to whether a point can be eliminated from the data set. In general, the researcher required that there be a reason (e.g., some factual anomaly) to justify data elimination. In cases where one or more outliers had significant influence on the model but no reason existed to believe that the outlier was caused by invalid data, two separate analyses were conducted, one including outliers and one excluding outliers.

The survey results include a detailed analysis of the demographics of the respondents. This information relates to the respondents, the types of systems being proposed, factual details of the proposal efforts and information about the level of confidence respondents have in their results. Appendix D provides a more detailed assessment of the parameters and demographics of this analysis.

The significance of factors is analyzed using univariate statistical analysis techniques. Each factor is analyzed for the level of statistical correlation with the probability of contract award. There are two types of factors to be considered, categorical factors and continuous factors. Categorical factors can only assume a finite number of levels (e.g. yes/no, levels 1-5). On the other hand, continuous variables assume values in an interval or in a collection of multiple intervals [Hines and Montgomery, 1972]. Generally, for categorical factors of contract awards, several statistical tests are performed to determine whether there is a high likelihood that the factors are correlated with the probability of contract award. These include analysis of variance, Pearson's Chi-square test and an exact version of Pearson's Chi-square test [IBM, 2013]. If any of the tests reveals an 80% or higher likelihood that the factor is correlated with an increased probability of contract award, then various post-hoc multiple comparison tests (e.g., Scheffé and Tukey) [Dean and Voss, 1999] are performed against various groups of data based upon the results.

For data collected where factors are represented on a continuous scale (e.g. ratio of level of investment for defining the proposed offering in the proposal to the number of estimated engineering labor hours on the development portion of the contract if the proposal is captured), a logistic regression model is used to attempt to fit the data. Binary logistic regression models are discussed in Section 2.2.3.2. Bootstrapping is used to determine whether tests designed for large samples of data (e.g., Wald test) yield valid results. If bootstrapping confirms the appropriateness of large sample tests, a Wald test [Kutner et al., 2005] is used to determine how likely it is that the factor analyzed is actually correlated with an increased probability of contract award. If not, statistics derived from bootstrapping as well as alternative measures such as profile-likelihood confidence intervals centered around Firth's estimate [Firth, 1993] characterize the significance of the relationship.

After factors have been analyzed in isolation, a number of approaches are considered to develop a *P* function of multiple independent variables from a subset of the factors analyzed. Various techniques for model building are used including best subsets method and stepwise approaches [Kutner et al., 2005; Montgomery et al., 2012]. A number of metrics, including Akaike's Information Criteria (AIC) and Schwartz Criteria (SC) [Kutner et al., 2005] can be used to compare candidate models and select a best model. Another important metric by which to evaluate a multivariate logistic regression model is events per variable (EPV). The complexity of a logistic regression model is fundamentally constrained by the data that is used to develop the model. The more valid data collected, the more parameters can validly be included in a regression model. A number of researchers address the issue of what EPV is required. An event is defined as the less-frequent (e.g., non-awards) of the two possible binary outcomes [Peduzzi et al., 1996]. This is by definition, less than or equal to half the total number of valid observations. A rule of thumb for a stable logistic regression model is that one needs a minimum of 10 events per variable [Harrell et al., 1985]. What is meant by variable is a parameter in the regression model, and the intercept term counts. Categorical variables with more than two values actually require more than one parameter. If there are *n* levels for the categorical variable, then there is a need for n-1 parameters. By this standard, to have a model with one continuous independent variable or one binary categorical variable would require a minimum of 20 proposals where a contract was awarded and 20 where a contract was not awarded. Appendix D provides additional detail related to the use of these multivariate techniques.

The next section describes the process of implementing a DSS based upon an analysis of the survey responses.

3.10 Implementing the Modeling Framework

The DSS implementation consists of a number of modules developed using Matlab and Matlab toolboxes. Ultimately, the DSS seeks to estimate the global optima over the feasible region. Mathwork's Global Optimization Toolbox is used to increase the probability that the estimated optimal solution is the true, global optimal solution within the feasible region. Ideally, the P function would be constructed using multivariate regression analysis using historical data. This research seeks to apply multivariate regression analysis to the survey results to derive model parameter values. However, the quantity of data collected by the survey was inadequate to support the derivation of a statistically stable objective function with sufficient complexity to address the research question. Because these requirements could not be met, a more general framework called the Systems Engineering Proposal Management Optimization Framework (SEPOMF) is introduced instead of the originally planned DSS. The SEPOMF provides guidance for how to develop a DSS to maximize the probability of contract award by optimally using systems engineering on a proposal of interest by leveraging historical data from past proposal efforts. The SEPOMF is flexible so that even if historical data has not been categorized exactly as is recommended by the particular activities and skill levels defined in this dissertation, there is still the potential to use that historical data. This creates the potential for organizations to leverage their already collected historical data.

To illustrate the SEPOMF, an example DSS was developed using the SEPOMF concept that allows the user to select between two sets of decision variables. For each set, there exists a *P* function with made up parameter values. These parameter values are arbitrarily defined by the researcher to demonstrate the potential usefulness of the SEPOMF in situations where adequate data exists. A detailed discussion of the DSS implementation can be found in Appendix E.

3.11 Validating the Modeling Framework

Once a DSS is developed, it must be validated. The original research design was to validate the DSS using the criteria defined in Table 2.6. As the survey data did not support the generation of a single data-driven model, the validation exercise instead focuses on the validity of the SEPOMF modeling framework. The model tests described in Table 2.6 were carefully analyzed for applicability to an optimization modeling framework. Based upon this analysis of the model tests, which is described in detail in Appendix G, a set of questions for validators to answer was derived. The validators were selected based upon knowledge of the use of systems engineering on proposals and also an understanding of technical topics relevant to the modeling framework. The criteria for validating the SEPOMF was that a validator had to have expertise in systems engineering, and understanding of regression analysis and optimization models, experience using systems engineering on proposals and experience with at least five different proposal efforts. The validators were also asked questions to verify that they met the criteria to validate the model. Validator inputs were analyzed to determine if there were any modifications required to the modeling framework or how the modeling framework is presented.

3.12 Study Limitations

The study exclusively examines the use of systems engineering on proposals in the context of the maximizing the probability of contract awards. Therefore, following the recommendations of this research may help organizations be awarded more contracts, but there is no particular reason to believe that following the recommendations will necessarily achieve other important objectives for stakeholders of proposals. In auction theory there is such a thing as the “Winner’s Curse” [Kagel and Levin, 1986]. The “Winner’s Curse” means that whoever wins an auction likely overpays for the prize being sought. There is an analogy to proposal management. It may be that proposed offerings that correspond to a very high probability of contract award offer such a favorable deal to the acquirer that the supplier organization cannot execute the contract profitably. In such a case, the use of systems engineering that favors a high probability of contract award may very well not favor making a profit.

Another limitation of the study is that it focuses on how to use systems engineering to optimize a particular proposal project. This may sub-optimize other objectives. Most organizations execute multiple projects simultaneously, some proposal projects and some projects to execute already captured proposals. In such a multi-project environment, decisions to optimize one project, in some cases, have been shown to negatively impact other projects over an extended period of time [Abdel-Hamid, 1993].

There are also limitations for which scenarios a DSS derived using the SEPOMF is directly applicable to. For example, a DSS assumes that attributes of the organization are fixed at the time resource allocation decisions are made on a proposal project. This may not always be the case. The researcher has been involved in a multi-year “must win” proposal effort where particular types of contributors were needed who were not available in the organization at the time the proposal effort began. When this was recognized, the organization committed to a recruitment and hiring campaign to attract the needed talent, and had the employees sufficiently familiar with the proposal effort by the time of submission.

The researcher is also aware of situations where organizations targeted particular contract opportunities that required the organization to hold a higher capability maturity rating than the organization held at the time that the decision was made to pursue the particular proposal opportunity. In such cases, the organization committed to advancing the maturity of their processes in parallel with defining a proposed offering for the proposal. In this case, the capability maturity level was increasing throughout the process of defining the proposed offering.

3.13 Threats to Validity

This section examines threats to validity, focusing on threats to the validity of the findings gleaned from the survey research. Threats to validity can be internal or external. Findings from a study have internal validity if effects observed in the dependent variable are actually caused by the independent variable and not by other factors [Rubin and Babbie, 2007]. External validity refers to whether the causal relationships from the study can be generalized beyond the study conditions [Rubin and Babbie, 2007]. In this research, as with all research using surveys to collect data, there are threats to validity. Table 3.6 provides a description of a number of different internal and external threats to validity of the survey research as well as any actions taken to mitigate each threat. Threats to validity to using decision support systems developed using the SEPOMF are discussed in Appendix E.

Table 3.6 Overview of Threats to Validity for Survey Research

Name	Internal/ External	Description	Status for Survey	Mitigations
Non-response	Internal	Some people targeted for the survey do not respond, and those who do not respond may have systematically different responses than those who do respond.	Less than 1% of the people targeted responded. However, subjects were targeted based upon their membership in professional organizations. It is expected that many of those people do not possess the requisite knowledge to respond.	Many people were solicited by multiple means, including direct emails and website posts.
Dropout	Internal	Respondents begin the survey but do not complete it.	For the Survey Monkey version of the survey, the dropout rate was over 98%.	Emails, website posts, and the first page of the survey clearly articulate the purpose, qualifications and time commitment.
Too Many Success Stories	Internal	Respondents may have been more likely to report proposals that resulted in contract awards than those that did not.	For the survey, 43 of the 62 valid responses reported contract awards. It is unknown whether this is representative.	Contact materials and the survey directions clearly stated that both proposals that resulted in contract awards and proposals that did not result in contract awards are of interest.
Failure to Read or Understand Directions	Internal	Respondents do not read directions or do not have adequate English-language reading skills to understand the directions.	Directions clearly indicate that only respondents who know the award status of their proposal effort should respond. Many of the respondents were from countries where English is not the primary language.	Directions were repeated numerous times in emails, website posts and in the survey itself.

Table 3.6—*Continued*

Undocumented Effort	Internal	Labor effort for activities to define the proposed offering is not accurately captured in the survey because they were never recorded. These labor hours may have been financed by another project (e.g., research and development) or they may have been uncompensated (e.g., people working over their standard 40 hours/week).	Several entries report a very small investment in systems engineering for the core activities for defining the proposed offering compared to the contract size (less than 0.5% of the total engineering hours) yet report contracts were awarded.	Tasks were described not in terms of accounting guidelines, but with respect to the activities undertaken. Respondents were instructed to provide their best estimates for various categories.
Representativeness of Sample	External	Members of the targeted professional societies may not be representative of the general population of people who must make resource allocation decisions for proposal efforts.	This is unknown. INCOSE includes participants from a number of different industries, but a large portion of INCOSE members work in the defense sector. NDIA is primarily defense.	No mitigations. Findings should be carefully analyzed for applicability if used in a context other than weapons systems or defense systems.

Chapter 4

Results

This chapter presents the results of the research. These results are analyzed to evaluate the detailed hypotheses described in Chapter 3. Then, the results and hypotheses conclusions are used to address the detailed research questions defined in Chapter 3.

4.1 Research Results

This section presents the results of the research. This includes the results of administering the survey, results of the survey analysis, implications of the analysis of the survey data for developing a valid model, results of implementing the modeling framework and results of validating the modeling framework.

4.1.1 Results of Administering the Survey

This section describes how the survey was administered and discusses key demographics. This information provides additional context to help readers evaluate the applicability of the survey analysis results and findings to their proposal scenarios. This section provides highlights related to administering the survey. For a more detailed and complete set of information about the survey, see Appendix D.

4.1.1.1 Survey Administration Details

The data used in the analysis came from three collections: field testing, a first collection and a second collection. While field tests were conducted primarily to validate the survey instrument, the decision was made to use inputs from the field tests as data for the survey analysis. This was because the field test inputs reported actual proposal efforts that were within the scope of analysis. Field testing occurred between August 5, 2012 and August 30, 2012. There were two versions of the survey coded using different software and available on different locations on the internet. A second version was required to overcome the technical issues with the first version that are discussed in Section 3.7.3. The first version was used to collect data from September 21, 2012 to October 15, 2012. The second version was used to collect data from November 13, 2012 to December 15, 2012.

4.1.1.2 Demographic Information

This section describes the demographics of the survey respondents. Demographics can provide useful contextual information for evaluating how applicable the survey analysis results are to a particular proposal opportunity. The demographics presented pertain exclusively to valid responses. Appendix D includes a discussion about the results of applying the validation checks that addresses the number of invalid responses and reasons why various responses were not deemed valid. Appendix D also provides a more extensive discussion of the data presented in this section and provides additional demographic information and statistics beyond what is included here (e.g., size of proposal team, location of customer's home office).

Figure 4.1 describes the distribution of responses between proposals where a contract was awarded and proposals where no contract was awarded. There were 43 contract awards and 19 non-awards of the 62 valid responses.

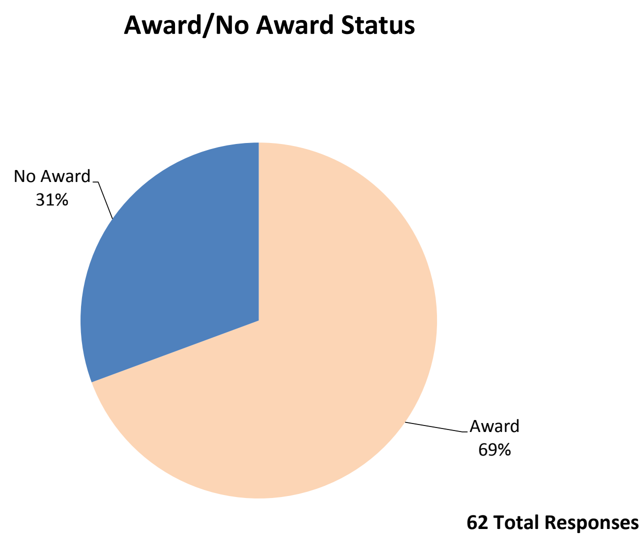


Figure 4.1 Percentage of Proposals Resulting in Contract Award/No Contract Award

Respondents were asked to specify whether they are members of the International Council on Systems Engineering (INCOSE), the National Defense Industrial Association (NDIA) or the Project Management Institute (PMI). This question was asked primarily so that it would be clear how extensible the results of the factor analysis are to the systems engineering community in general and to gauge how effective efforts were to promote the survey within various communities. 92% of respondents reported being INCOSE members, 20% reported being NDIA members and 24% reported being PMI members. All of the respondents selected at least one of the three professional associations.

Figure 4.2 describes the respondents' experience by role. Almost 90% of the respondents reported backgrounds of serving as individual contributors in systems engineering, and over 80% had a background in other engineering disciplines. More than 70% of respondents had served as a lead systems engineer. More than 60% had been project managers, and more than 50% had been lead engineers or functional managers in charge of organizations responsible for systems engineering.

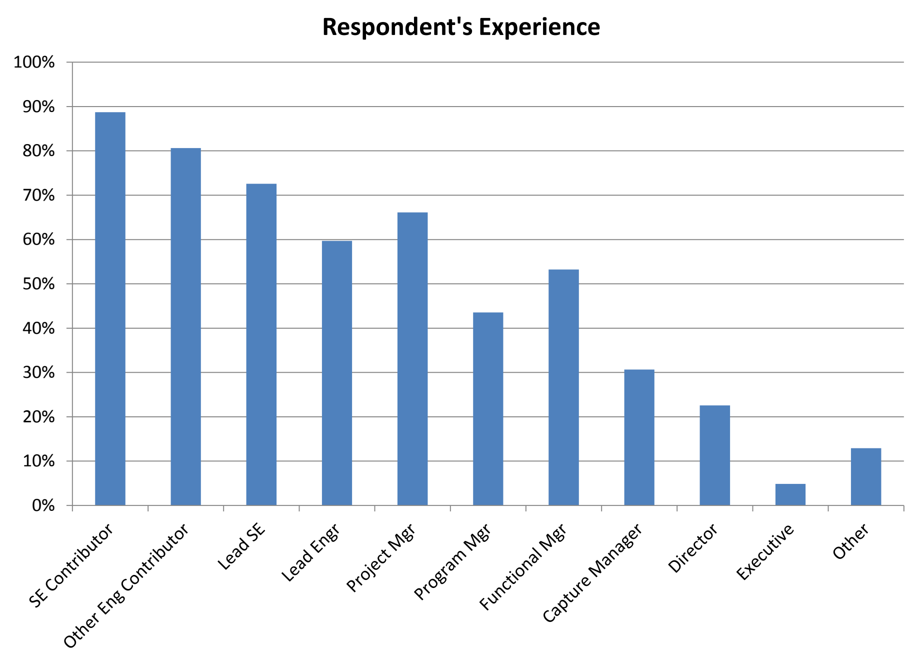


Figure 4.2 Respondents' Experience by Role

Figure 4.3 displays the number of proposal per sector. The sectors include commercial, military/defense and government non-military. Over 50% of the proposals report contracts for military/defense projects. Figure 4.4 reports a complementary finding: almost half of the proposals were for weapons systems.

Many of the survey questions pertain to the prime contractor and the relationships between the prime contractor and customer. 81% of respondents were from prime contractors while 19% were from subcontractors.

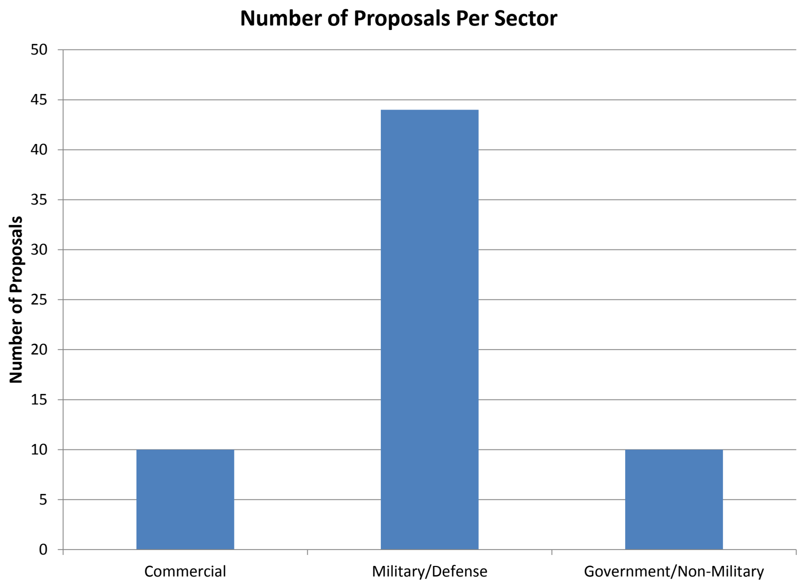


Figure 4.3 Number of Proposals per Sector

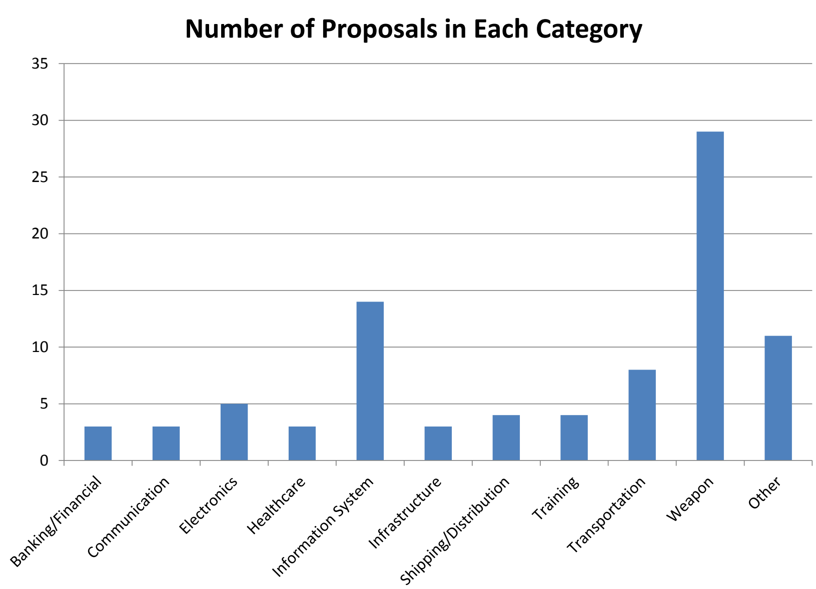


Figure 4.4 Number of Proposals Selecting Each Category

A final demographic question respondents were asked was to indicate their level of confidence in their responses, given choices from “not confident at all” to “extremely confident”.

94% of respondents indicated that they were confident, very confident or extremely confident about their responses to qualitative items. 86% of respondents indicated that they were confident, very confident or extremely confident about their responses to quantitative items. No respondents reported not being confident at all in their responses.

4.1.2 Results of Performing Factor Analysis

This section discusses the results of the factor analysis, focusing on highlights from the analysis. This section includes the results from three relevant statistical tests for categorical factors and post-hoc multiple comparisons, the results from two tests for continuous predictors and discusses attempts to develop a multivariate logistic regression function from the data. For a more detailed and complete discussion of the analysis methodology, see Appendix D.

4.1.2.1 Results from Single Factor Analysis

Examining each factor individually allows the reader to understand if a particular factor appears to be correlated with an increased probability of contract award. Table 4.1 describes the results of three relevant statistical tests for categorical factors: Analysis of Variance (ANOVA), Pearson's Chi-square test and an exact version of Pearson's Chi-square test. The values that are displayed are ρ-values. ρ-values are the smallest value of alpha (i.e., significance level), that would allow the null hypothesis to be rejected [Dean and Voss, 1999]. These values are between 0 and 1 by definition. For the subsequent discussions in this dissertation, ρ-values of less than 0.05 signify strongly significant relationships and ρ-values between 0.05 and 0.2 signify mildly significant relationships. Rows are greyed for factors for which at least one of the tests estimated a ρ-value below 0.2.

The factors with potential significance are (1) type of contract, (2) system and software process maturity of prime, (3) level of project manager experience and (4) level of customer satisfaction on previous or ongoing effort. For these four factors, further post-hoc analysis was performed to explore the nature of the relationship and see if any more specific statements could be made. For each of these factors, one contrast was defined that was suggested by the data to clarify differences between factor levels. Both Sheffe and Tukey [Dean and Voss, 1999] methods were used to assess the significance of that contrast. The details of this post-hoc analysis can be found in Appendix D. The contrasts related to type of contract and to level of customer satisfaction on previous or ongoing effort both proved significant. As a result of this analysis, the following statements are made:

1.) Type of contract - Proposals for contracts that are firm fixed price (FFP) have a lower probability of being awarded contracts than proposals for contracts with arrangements other than FFP.

2.) Level of customer satisfaction on previous or ongoing effort - Proposals submitted to customers that are assessed by the supplier to be very satisfied with previous or ongoing contract work from the supplier are more likely to result in contract awards than proposals submitted to customers who are assessed by the supplier to be less than very satisfied with previous or ongoing contract work from the supplier.

Table 4.1 Statistical Results Overview - Categorical Factors

Factor	ANOVA	Pearson	Exact
Type of Contact	0.086	0.086	0.088
Motive - Profit from Development	0.826	0.822	1.000
Motive - Profit After Development	0.359	0.351	0.410
Motive - Future Work	0.376	0.368	0.545
Motive - Future Relationships	0.851	0.848	1.000
Motive - Satisfy Existing Customer	0.716	0.710	0.784
Motive - Block Competitor	0.327	0.319	0.422
Systems and Software Process Maturity of Respondent Organization	0.501	0.481	0.515
System and Software Process Maturity on Project	0.688	0.668	0.727
System and Software Process Maturity of Prime	0.174	0.169	0.169
Level of Access	0.637	0.620	0.692
Level of Project Manager Experience	0.122	0.122	0.101
Rank - Market Share	0.428	0.409	0.437
Rank - Prestige	0.336	0.321	0.330
Extent of Relationship with Proposal Customer	0.524	0.507	0.520
Level of Customer Satisfaction on Previous or Ongoing Effort	0.037	0.042	0.011

No significant contrasts were found related to system and software process maturity of the prime or the level of experience of the proposal project manager. Because of the relatively small number of proposal efforts reported, ρ-values above 0.2 do not necessarily mean that the factor is insignificant. For such factors, it is inconclusive whether or not the factor was significant because of the limited quantity of survey responses. It is possible that with a larger or more homogenous data set some of these factors very well may be determined to be significant.

Table 4.2 displays the estimated ρ-values for variables relating to the number of contacts with the customer during the proposal effort normalized by the total number of engineering labor hours to be worked on the development portion of the contract. Based upon bootstrap testing, traditional sorts of asymptotic tests for significance such as Wald tests were found to be inapplicable because the sample size (i.e., number of surveys) was too small. Therefore, ρ-values from bootstrapping and profile likelihood confidence intervals centered on Firth's method [Firth, 1993] are used. Just as in Table 4.1, rows corresponding to factors for which at least one test found a ρ-value of less than 0.2 are greyed. The rows that are greyed are the number of video conference contacts and the number of in-person contacts. For most of the proposal efforts reported, the total number of contacts was largely comprised of written contacts. Because of this heavy weighting of written contacts in the total contacts statistics, the total number of contacts does not appear significant.

Table 4.2 Statistical Results Overview - Contacts with the Customer (Normalized)

Contact Type	ρ-value Bootstrap	ρ-value Firth
Number of written contacts	0.496	0.473
Number of telephone contacts	0.286	0.353
Number of video conference contacts	$0.05 < p < 0.1$	0.341
Number of in-person contacts	0.064	0.289
Total number of contacts	0.328	0.404

Table 4.3 displays the estimated ρ-values for what is labeled DPORatio. DPORatio is defined throughout this dissertation as the number of systems engineering labor hours spent defining the proposed offering on the proposal divided by the total number of engineering labor hours to be spent on the development portion of the contract. Some portions of this, related to certain activities or particular skill levels, are shown with a dash after DPORatio. For example,

the number of systems engineering labor hours spent on source requirements divided by the total number of engineering labor hours to be spent on the development portion of the contract, is referred to as DPORatio – source requirements. The same two methods were used to derive *ρ*-values as were used in Table 4.2. DPORatio was found to be mildly significant. With the exception of source requirements, the level of effort devoted to the other four systems engineering activities (system technical requirements, architecture, cost estimation and decision analysis) appeared significant or mildly significant. The level of effort corresponding to all of the four skill levels appears significant or mildly significant.

Table 4.3 Statistical Results Overview - Level of Effort on Proposals (Normalized)

Numerator	*ρ*-value Bootstrap	*ρ*-value Firth
DPORatio - all skill levels	0.052	0.060
By Activity (all Skill Levels)		
DPORatio - source requirements	0.271	0.275
DPORatio - system technical requirements	0.045	0.066
DPORatio - architecture	0.044	0.154
DPORatio - cost estimation	0.064	0.167
DPORatio - decision analysis	0.213	0.195
By Skill Level (all Activities)		
DPORatio - beginner skill levels	$0.1 < p < 0.2$	0.173
DPORatio – intermediate skill levels	0.127	0.233
DPORatio - advanced skill levels	0.087	0.144
DPORatio - expert skill levels	0.103	0.199

A major limitation of examining factors individually is that the relationships between the factors are not accounted for. In analyses such as this, sometimes multiple important variables will have strong statistical correlations. Further analysis may determine that one of the variables is important and the rest are just correlated with the important variable. Before focusing on optimization, multiple factors need to be considered simultaneously. The next section presents such an analysis.

4.1.2.2 Results of Multivariate Statistical Analysis

This section will discuss the attempts that were made to develop a multivariate logistic regression model with the available data. As has been mentioned in both Chapter 1 and Chapter 3, this analysis does not result in a stable, valid model that includes all of the factors believed to be important. In the process of attempting to develop such a model, a number of observations are made regarding the survey data. Because of this, these attempts and important observations along the way are addressed.

Because the outcome of each case is binary (a contract was awarded or no contract was awarded) and because the model is designed to estimate the probability of achieving a particular outcome (i.e., a contract award), a binary logistic regression model with multiple independent variables is used. Categorical and continuous factors are considered simultaneously. By focusing on multiple factors all at once, a model can be developed where the level of correlation between the factors included in the model is minimal.

Even though there were 62 valid responses to the survey, only 37 of those were sufficiently complete to be used to develop the multivariate logistic regression model. Of the 37, there were 22 where contracts were awarded and 15 where no contracts were awarded. Based upon the standard of EPV of 10 or greater (as discussed in Section 3.9), the stability of models that consider single continuous or binary independent variables have questionable stability. As more variables are considered, the EPV goes down and the likely stability of the model goes down. Nonetheless, experimentation was performed where attempts were made to develop multivariate models.

The following set of factors is considered in the model selection process: type of contract, system and software process maturity of prime, level of project manager experience, and level of customer satisfaction on previous or ongoing effort. Also, DPORatio was considered, number of video conferences divided by total number of engineering hours to be worked on the development portion of the contract was considered, and number of in-person contacts divided by total number of engineering hours to be worked on the development portion of the contract was considered. A number of automated model selection methods were applied to generate candidate models and a number of standard metrics were used to compare the various candidate models including AIC, SC and EPV. This is documented in Appendix D.

After extensive analysis was performed, the model that appeared to be the best multivariate model included type of contract and DPORatio. After a more careful analysis was performed, all of the proposals reported (within the set of 37 complete responses) that reported cost plus percentage of cost or cost plus fixed fee were for cases where a contract was awarded. Therefore, the model predicted over 80% probability of contract award for these types of contracts even if no investment was made in defining the proposed offering. This prediction is believed to be an artifact of the particular data set used and not something that can be generalized. With this data, there does not appear to be a believable model that can be built to predict the probability of contract award for cost plus type contracts. However, there is adequate data to address FFP contracts somewhat more robustly.

There are 24 responses that reported FFP contracts, 11 of which correspond to contract awards and 13 that correspond to non-awards. When just these data points are used, a model can be developed with an EPV of 5.5. This model and the data used to develop this model are depicted in Figure 4.5. Figure 4.5 displays the values for DPORatio of all the FFP contracts. Each contract is represented by a circle. The circles with a y value of 0 correspond to non-awards and the circles with a y value of 1 correspond to contract awards. The curve that passes through (0, 0.35) and (100, 0.78) is the fitted model. Depending on the metric used, the probability that the relationship between DPORatio and the probability of contract award is significant is somewhere between 83% and 95%. A careful examination of Figure 4.5 reveals that several of the data points appear to be outliers. These all correspond to contracts with very high values for DPORatio. Both inspection and an exploratory statistical technique known as boxplot analysis concur that the three data points with the largest DPORatio are all outliers. It is useful to re-examine this relationship removing these three outliers. The same analysis with the outliers removed is depicted in Figure 4.6. The scale on the x-axis for Figure 4.6 is not as wide as in Figure 4.5 because the three largest data points have been removed. The same symbols are used for the same purposes as were used in Figure 4.5. With fewer data points, the statistical confidence in the significance of the relationship drops to between 75% and 80%

confident that the relationship is significant. There is no reason to necessarily believe that the statistics derived from the data set with the outliers removed are necessarily more or less representative of the true relationship than the statistics derived from the data set including the outliers. However, there is the potential that there is something fundamentally different about these three points.

4.1.2.3 Implications of Analysis of Survey Data for Developing a Valid Model

Given the limited quantity of data as well as the data characteristics, a stable model with adequate complexity could not be developed from the survey data. Without any other sources of proposal data to use, the decision was made to present the framework for developing and solving optimization models using a set of available data. It is valuable to develop and present this concept because it is expected that within some organizations, there may exist or may soon exist a sufficient quantity of historical proposal efforts to apply this technique to. Developing example models to illustrate this framework will also help clarify how historical proposal data can be used. This may help organizations who do not have an established process in place for recording proposal effort data to begin collecting data in a smart way that can later be used for the benefit of the organization.

4.1.3 Results of Implementing the Modeling Framework

The Systems Engineering Proposal Optimization Modeling Framework (SEPOMF) provides guidance for how to develop a DSS to maximize the probability of contract award by leveraging historical data from past proposal efforts. The SEPOMF is flexible so that even if historical data has not been categorized exactly as is recommended by the particular activities and skill levels defined in this dissertation, there is still the potential to use that historical data. This provides the option for organizations to leverage their already collected historical data.

The SEPOMF includes a formal definition of the optimization problem that must be solved by each SEPOMF DSS. The SEPOMF also provides guidance for decision makers through the complete process of identifying decision variables, determining what historical data is usable, deriving and selecting an objective function from the historical data and defining and solving an optimization problem. More details related to this can be found in Appendix E.

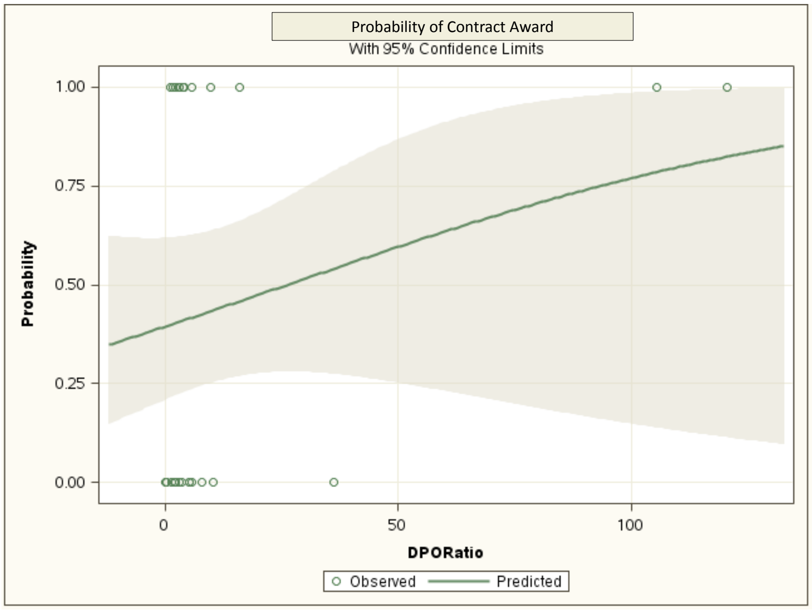


Figure 4.5 *P* vs. DPORatio - Firm Fixed Price Contracts

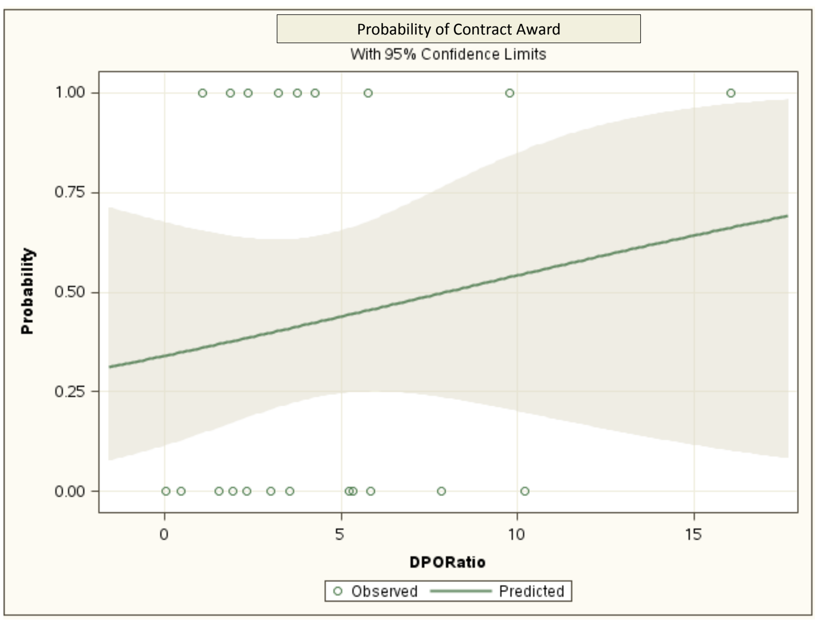


Figure 4.6 *P* vs. DPORatio - Firm Fixed Price Contracts

Three examples of the modeling framework that are developed for illustrative purposes are provided in Appendix F. While the SEPOMF is capable of considering both systems engineering activities and employee skill levels simultaneously, the number of variables and the number of required model parameters tends to become large quickly when this is done, and this is only feasible when there are hundreds or even thousands of historical proposal efforts to develop *P* functions with. Therefore, the examples provided in Appendix F each focus either on allocating budget between systems engineering activities (i.e., Example #1 and Example #2) or allocating budget between employees with varying skill levels and varying labor rates (i.e., Example #3).

The primary means for expressing the implementation of the SEPOMF is to examine inputs and outputs of decision support systems developed using the SEPOMF. All displays of inputs and some displays of outputs are taken from screen captures from the portion of the example DSS that depicts the dialogue between user and DSS. Other outputs are taken from graphics generated by the DSS. Because the objective function used for these example decision support systems involves parameters that are made up for illustrative purposes, the recommended budget allocations from the examples should not be used by decision makers to allocate their resources. Rather, the methodological example should be followed, and decision makers should use historical data from within their organizations to develop the objective function. Appendix E provides significant explanation and guidance for how to actually use the SEPOMF to develop a DSS to optimize the use of systems engineering on proposals using real data from a particular organization.

The example decision support systems were developed solely for illustrative purposes and are of interest primarily to people seeking to develop similar systems. Because of this, details are shown only in Appendix F. Table 4.4 provides a legend to the inputs and outputs for the various example decision support systems to help readers navigate the information in Appendix F.

Table 4.4 Legend to Inputs and Outputs for Example Decision Support Systems

	DSS #1	DSS #2	DSS #3
Factors included	• System Technical Requirements • Architecture • Cost Estimation	• Source Requirements • System Technical Requirements • Architecture • Cost Estimation • Decision Analysis	• Beginner • Intermediate • Advanced • Expert
Inputs	Figure F.1 Figure F.2 Figure F.3	Figure F.6	Figure F.9
Outputs	Figure F.4 Figure F.5	Figure F.7 Figure F.8	Figure F.10 Figure F.11

4.1.4 Results of Validating the Modeling Framework

The SEPOMF was validated by four different experts. The four validators have experience in four different organizations, and all four validators met the criteria to validate discussed in Section 3.11. Two validators have industry experience from the same organization, but one validator's industry experience spans two organizations. Interviews with the validators occurred between July 30, 2013 and August 17, 2013. All four validators provided written feedback. Written inputs were received between August 8, 2013 and September 3, 2013. Appendix G provides a detailed discussion of this validation process, the written inputs provided by each validator and the resolutions to address the inputs of the validators.

Resolutions to validator comments were formulated by comparing the different validator responses to see if common issues were raised by more than one validator, by evaluating the relevance of each comment with respect to the larger research goals of the study and by determining if each issue relates to a fundamental shortcoming of the research or is simply caused by the limited subset of information about the research that was provided to validators at the time of validation. In some cases, resolutions describe modifications to this dissertation or the SEPOMF model definition. In some instances, the resolutions highlight particular features of the SEPOMF that are believed to address the validators concerns but may not have been clearly communicated by the materials given validators to review. In other cases, the resolutions are simply responses to the validator comments.

Of primary interest are the resolutions that resulted in a change to the model or presentation of information in the dissertation. As a result of analyzing the feedback from validators (see Table G.7), the following significant changes were made:

1.) Section 1.9 has been added to the introduction chapter to clarify that this dissertation is written to an audience with significant technical depth in analytical techniques as well as an audience with a thorough understanding of the core systems engineering and proposal management concepts involved.

2.) To clarify assumptions and various conditions that need to be in place before using the decision support systems developed using the SEPOMF, additional preconditions have been added to the formal use case definition in Table 3.5.

3.) A discussion has been added in Section 3.5.3 to help the reader understand the mathematical importance of the constraints u and t to the optimization process. This discussion addresses how constraints can limit the feasible region and potentially exclude otherwise optimal solutions.

4.) A discussion has been added in Section 3.5.3 to explain the mathematical behavior of a logistic regression model and why this mathematical behavior makes a logistic regression model particularly well suited for an objective function for the optimization problem at the core of the decision support systems developed using the SEPOMF.

5.) The need to reconsider risk as a potential factor in future research is added to the discussion of future work in Section 5.4.

In addition to actual changes, several themes persisted in the validator feedback that require a mention. The following summarizes these key themes:

1.) Collecting sufficient data to have an example DSS derived from actual data and developed using the SEPOMF will aid in user community acceptance of this modeling concept.

Until this is done, practitioners may view the SEPOMF concept as largely theoretical and unproven.

2.) When presenting the SEPOMF, care must be taken to describe the timeframe in the system lifecycle and proposal process where the optimization is occurring and what the preconditions for use are.

3.) It needs to be emphasized when presenting the SEPOMF that the framework is flexible as to what variables are included other than the decision variables. The set of variables explored in the survey is a starting point, but likely not the definitive set. Further research should be done regarding what these variables are.

4.2 Detailed Hypotheses Conclusions

Table 4.5 revisits the hypotheses formulated in Table 3.3 and draws conclusions as to whether the data supports each hypothesis or not. In many cases, the results are inconclusive. The first column is the hypothesis identifier, the second column is the statement of the hypothesis and the third column is the test results for the hypothesis.

In order to make these assessments, formal hypothesis testing was performed. The formalism of hypothesis testing (e.g., null hypothesis, alternative hypothesis) is not adopted in this presentation because it can be confusing to uninitiated readers. In formal hypothesis testing, the stated hypotheses are the alternate hypotheses and the null hypotheses described the complement to the intended relationship. In some cases, the null hypothesis can be rejected. In these cases, the alternate hypotheses are accepted. For such cases, test results reported that the hypothesis is supported by the data. In other cases, the null hypothesis failed to be rejected but there was not enough data to accept the null hypothesis. In such cases, the test results for the hypothesis tests are inconclusive.

4.3 Answers to Detailed Research Questions

This section revisits the detailed questions posed in Section 3.1 and provides a response to each of these questions based upon the results of the research.

Table 4.5 Conclusions of Detailed Hypotheses

Hypothesis Number	Hypothesis Statement	Test Results for Hypothesis
1-A	As the ratio of number of systems engineering labor hours defining the proposed offering to the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	Mildly supported by the data – Depending upon what metric is used and whether outliers are considered, there is between a 75% and 95% likelihood that this relationship exists.
2-A	As the number of written interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	Inconclusive - The data does not show a statistically significant relationship between an increased relative number of written interactions between the prime contractor and the customer and an increased probability of contract award. However, there is an insufficient quantity of data points to demonstrate that this factor is not significant.
2-B	As the number of telephone interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	Inconclusive - The data does not show a statistically significant relationship between an increased relative number of telephone interactions between the prime contractor and the customer and an increased probability of contract award. However, there is an insufficient quantity of data to demonstrate that this factor is not significant.
2-C	As the number of video conference interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	Mildly supported by the data – By some metrics, there is greater than a 90% chance that as the relative number of video conference contacts increases the probability of contract award also increases. By other metrics, no statistically significant relationship exists.
2-D	As the number of in-person interactions between the prime contractor and the customer divided by the total number of engineering labor hours to be worked on the development portion of the contract increases, the probability of contract award increases.	Mildly supported by the data - By some metrics, there is greater than a 90% chance that this relationship exists. By other metrics, no statistically significant relationship exists.
2-E	As the level of systems and software process maturity of the prime contractor increases, the probability of contract award increases.	Inconclusive – The data does not show that level of system and software process maturity of the prime is a significant factor. However, there is an insufficient quantity of data to demonstrate that this factor is not significant.
2-F	As the perceived level of customer satisfaction with previous	Strongly supported by the data – By all metrics considered,

Table 4.5—*Continued*

	or ongoing contract efforts increases, the probability of contract award increases.	there is over a 95% chance that the perceived level of customer satisfaction is a statistically significant factor. Post hoc multiple comparison analysis indicates that proposals where the customer is very satisfied with previous or ongoing contract work were significantly more likely to result in contract awards than proposals for which the customer was less than very satisfied.
2-G	As the competitive rank of the prime contractor with respect to market share increases, the probability of contract award increases.	Inconclusive – The data does not show that the competitive rank of the prime with respect to market share is a significant factor. However, there is an insufficient quantity of data to demonstrate that this factor is not significant.
2-H	As the competitive rank of the prime contractor with respect to prestige increases, the probability of contract award increases.	Inconclusive – The data does not show that the competitive rank of the prime with respect to prestige is a significant factor. However, there is an insufficient quantity of data to demonstrate that this factor is not significant.
2-I	As the experience level of the project manager of the proposal project for the prime contractor increases, the probability of contract award increases.	Inconclusive – The data did not show that level of project manager experience of the prime was a significant factor. However, there was an insufficient quantity of data to demonstrate that this factor is not significant.

4.3.1 Systems Engineering Factors and Proposal Success

Question: "What systems engineering related factors impact the probability of contract awards?"

Response: There are several factors that when analyzed individually have statistically significant correlations (see Section 4.1.2.1) with an increased probability of contract award. These are: (1) level of customer satisfaction with previous or ongoing contract efforts, (2) type of contract, (3) number of labor hours worked on the proposal performing systems engineering activities related to defining the proposed offering normalized by contract size, (4) number of video conference interactions with the customer normalized by contract size and (5) number of in-person interactions with the customer normalized by contract size.

Further analysis indicates that:

1.) Proposals for firm fixed priced contracts are less likely to result in contract awards than proposals for contracts with other type of contractual arrangements.

2.) Proposals where the customer is assessed by the supplier to be very satisfied with previous or ongoing contract work have a higher probability of contract award than proposals for customers who were assessed by the supplier to be less than very satisfied with previous or ongoing contract work.

3.) As the number of labor hours worked on the proposal performing systems engineering activities normalized by contract size increases, so does the probability of contract award.

4.) As the number of video conference interactions with customers normalized by contract size increases, so does the probability of contract award.

5.) As the number of in-person interactions with customers normalized by contract size increases, so does the probability of contract award.

One important consideration is that factors that are not found to be statistically correlated with an increased probability of contract award may in fact be demonstrated to be significant if a larger sample of data was available to analyze. The factors for which this is the case include: (1) level of system and software process maturity, (2) motivation for submitting the proposal, (3) number of interactions with the customer normalized by contract size, (4) competitive rank of the organization and (5) experience level of proposal project leadership. The factors selected for the survey were carefully chosen, but are likely not the definitive set of factors. There are other potentially important factors. In fact, validators identified a number of additional factors, and these are captured in the validator responses in Table G.3 through Table G.6. As these additional factors were identified after the survey, no data exists to substantiate or refute their significance. Nonetheless, some of these should be considered in future research.

The survey data does not support an analysis that properly accounts for the correlations and interactions between these various factors. The relatively small quantity of data and the fact that there are very few responses prohibits developing stable models that consider multiple factors simultaneously. Without being able to consider multiple factors simultaneously, it is impossible to identify which variables are truly important. It may be that key interactions of factors are very strongly correlated with an increased probability of contract award even if none of the constituent factors are strongly correlated by themselves.

4.3.2 Optimal Total Level of Engineering Effort

Question: "When defining the proposed offering to be included in a proposal, how does the ratio of *total labor hours worked performing systems engineering activities defining the proposed offering on the proposal effort* to *the estimated total number of engineering labor hours to be worked on development portion of the contract* impact the probability of being awarded a contract?"

Response: As the ratio of total labor hours worked performing systems engineering activities defining the proposed offering on the proposal effort to the estimated total number of engineering labor hours to be worked on development portion of the contract increases, the probability of contract award statistically increases. Estimates for the strength of this statistical relationship depend upon what metrics are used to estimate significance and how outliers are treated. As discussed in the response to the previous question causation has not been proven. Therefore, this does not necessarily imply that spending more labor hours on systems engineering activities related to defining the proposed offering for the proposal will actually cause the probability of contract award to increase.

4.3.3 Relative Allocation of Labor Hours for Various Systems Engineering Activities

Question: “When defining the proposed offering to be included in a proposal, what is the optimal percentage of the systems engineering labor hours to expend on each of the various SE activities (defining and validating source requirements, system technical requirements, architecture, cost estimation and decision analysis) to maximize the probability of being awarded a contract?"

Response: The SEPOMF modeling framework introduced provides a potential way for decision makers to optimize how they distribute labor hours across these various activities. This distribution of course depends upon the historical data used to calibrate the SEPOMF model as well as other important parameters, such as hourly rates for individuals performing each activity and the constraints on the total number of available labor hours to perform an activity. While the original design of this research was to leverage the survey data to derive an example DSS that will generate an optimal solution, not enough survey data was collected to accomplish this goal. Therefore, there does not exist an example optimal solution that can be generalized to help guide decision makers.

Based upon analysis and input from subject matter experts received during validation, there are several properties that optimal solutions will likely exhibit. Optimal solutions will likely involve some labor hours devoted to each of the systems engineering activities identified, especially systems technical requirements, architecture and cost estimation. These three activities were statistically correlated with an increased probability of contract award. Inputs from subject matter experts during validation indicate that a convincing business case must involve a solid understanding of the requirements and propose an architecture (including key technologies) that addresses those requirements with an acceptable level of risk. A major part of this risk assessment involves estimating costs. Another property optimal solutions will likely exhibit is that the percentage of labor hours to be devoted to various systems engineering activities may vary as the total budget for these activities increases or decreases. A third likely property of optimal solutions is that the recommended distribution of labor hours may depend upon the employee labor rates for these various activities. If employee labor to perform one of these activities costs more per hour than employee labor to perform another activity, the optimal distribution of labor hours across activities may be affected.

4.3.4 Relative Allocation of Labor Hours Across Various Skill Levels for Each Activity

A number of questions were posed relating to the optimal allocation of labor hours to employees with varying skill levels for various systems engineering activities. The same response is offered to address all of these questions. Each of the questions is restated here and then the response is provided.

1.) "When defining and validating source requirements in support of defining the proposed offering for a proposal, given labor costs by skill levels, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating source requirements?"

2.) "When defining and validating system technical requirements in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating system technical requirements?"

3.) "When defining and validating system architecture concepts in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in defining and validating system architecture concepts?"

4.) "When estimating costs of various candidate system solution concepts in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in cost estimation?"

5.) "When performing decision analysis to select a final solution in support of defining the proposed offering for a proposal, given labor costs by skill level, what is the optimal percentage breakdown of the systems engineering effort for contributors with beginner, intermediate, advanced and expert skill levels in performing decision analysis?"

Response: As with the optimal distribution of labor hours across the various systems engineering activities, the SEPOMF also provides a potential way to distribute labor hours across different contributors with varying skill levels. Like with the distribution of labor hours across various activities, the results from the survey data were not adequate to build a stable, statistically significant model. Therefore, organizations will have to collect enough data to develop these models before applying the SEPOMF.

There are several properties these optimal solutions will likely exhibit. First, the recommended distribution of employee skill levels will likely depend heavily on the relative differences in employee hourly labor costs. It is hypothesized that employees with more advanced skill levels are more effective at increasing the probability of contract awards per labor hour of their time than employees with less advanced skill levels. If this is true, the critical question is how much more effective. For large contracts, it may be that the optimal solution involves a mix of employees with different skill levels. This percentage mix may also depend upon the available budget. For smaller contracts where the absolute number of labor hours defining the proposed offering may be limited, the optimal solution may favor distributions where only an advanced or expert employee contributes for certain activities. While lower skill level, less expensive employees may be able to do routine work related to the proposals more efficiently, the additional costs of communications overhead [Abdel-Hamid and Madnick, 1991] may outweigh the seeming cost advantage of including the more junior employees.

Chapter 5

Conclusions and Future Work

The chapter begins with a summary of findings from the research. Recommendations are then offered for those who use systems engineering on proposals. Contributions of this research to both people using systems engineering on proposals and the scholarly literature are discussed. This chapter concludes with a discussion about future work.

5.1 Summary of Findings

This section explores the key findings that have emerged as a result of this dissertation research. This section first presents findings related to the factor analysis and then presents findings related to the optimization analysis.

5.1.1 Findings Related to Factor Analysis

The findings related to the factor analysis pertain to key relationships that were found to be significant in a statistical analysis of the survey results. The findings are:

1.) Proposal efforts for which the customer has been very satisfied with previous or ongoing contract work are more likely to result in contract awards then proposals for customers who are less than very satisfied.

2.) As the total number of labor hours worked on systems engineering activities relative to the contract size increases, the probability of contract award increases.

3.) As the number of labor hours worked defining and validating system technical requirements relative to the contract size increases, the probability of contract award increases.

4.) As the number of labor hours worked defining and validating the architecture relative to the contract size increases, the probability of contract award increases.

5.) As the number of labor hours worked on cost estimation relative to the contract size increases, the probability of contract award increases.

6.) As the number of labor hours worked on decision analysis relative to the contract size increases, the probability of contract award increases.

7.) As the number of face-to-face contacts relative to the contract size increases, the probability of contract award increases.

The analysis did not prove that any of the factors that are significant actually cause the probability of contract award to increase. Some factors that statistically correlate with an increased probability of contract award may just be correlated with other factors that cause the probability of contract award to increase. Because it is difficult to implement a controlled experiment related to the use of systems engineering on proposals, it may not be feasible to determine what factors truly cause an increase in the probability of contract award. Instead, statistical correlations and the wisdom of experts may have to be relied upon to drive decision making related to the use of systems engineering on proposals.

5.1.2 Findings Related to Optimization Analysis

This research has introduced a framework called the Systems Engineering Proposal Optimization Framework (SEPOMF) that guides an organization in optimizing the use of systems engineering on a proposal effort of interest by leveraging historical data. The framework addresses key activities including identifying the set of factors to consider, qualifying historical data to determine what subset to use, normalizing historical data by using information related to contract size, selecting an appropriate objective function relating the various factors to the probability of contract award, defining and solving an optimization problem using that objective function, and interpreting and acting on the results. The SEPOMF concept is sufficiently flexible that organizations can customize models to work with data that they may already have related to proposals and to allow different sets of factors to be considered depending on which factors are significant in what context. Because a stable SEPOMF model could not be developed with the survey data, an optimal strategy for allocating systems engineering resources on a proposal was not found.

5.2 Recommendations to Decision Makers

This section presents a number of recommended actions decision makers within an organization may take to more effectively use systems engineering on proposals. Due to the fact that the data really does not allow causation to be definitively proven, the recommendations that are made are derived from synthesizing the results of the statistical analysis with input from multiple experts and the author's professional experience. These expert opinions were elicited primarily during various validation exercises. These validation exercises included the validation of the set of factors, the validation of the survey questions, the field testing of the survey (i.e., validation of the survey instrument) and the validation of the SEPOMF. Based on this synthesis, the primary recommendations of this research are:

1.) Satisfy your existing customers – The survey data analysis, input from numerous subject matter experts who have experience as decision makers on complex systems and the author's professional experience all concur that customer satisfaction is a critical component of capturing new business. In fact, the statistical analysis showed a stronger relationship between the level of satisfaction of customers on previous or ongoing contract efforts with probability of contract award than it did between the budget spent on systems engineering on proposals normalized by contract size with the probability of contract award.

2.) Invest adequately in systems engineering activities while defining the proposed offering, focusing on key systems engineering activities – The analysis of the survey data confirms that spending more budget relative to contract size on system technical requirements definition and validation, architecture definition and validation, cost estimation and decision analysis is correlated with an increased probability of contract award. Of course, investing adequate resources in systems engineering defining the proposed offering is no guarantee that a viable system concept will emerge. There are cases where a lack of insight causes a significant investment to yield very little.

3.) Determine which factors are most critical to success – The survey data analysis found several statistically significant factors. However, the data set was not large enough to definitively rule out the potential significance of the other factors. It could be with a larger set of data factors that were not found to be significant in this analysis would in fact be found to be significant. In addition, a few promising factors were identified after the survey that should also be considered (see Appendix G). It is likely that with either a larger data set or a more homogenous data set taken exclusively from a single organization, additional factors will emerge as important.

4.) Use appropriately - As with any model, apply models developed using the SEPOMF with care because there is potential for misuse. Consider the context of the proposal opportunity and whether a model, such as one developed using the SEPOMF, makes sense before applying the SEPOMF. Also, models developed using the SEPOMF provide recommendations that are only as valid as the data that is used to calibrate the models. Even if the data used to calibrate is all good data, if there is something systematically different about the proposal effort of interest than the proposal efforts used to calibrate, the models can provide ill-advised recommendations. All recommendations by SEPOMF models should be carefully evaluated by experts to ensure that they make sense before resources are committed.

5.3 Contributions of this Research

This section focuses on the contributions of this research. The primary contributions of the research include a factor analysis study and a validated optimization modeling framework related to the use of systems engineering on proposals.

5.3.1 Factor Relationship Study

The primary contribution of this research is a factor relationship study that examines the statistical significance of the correlation between various systems engineering related factors and an increased probability of contract award. Included in this factor analysis is the identification of a few key factors that are significantly correlated with an increased probability of contract award such as type of contract, level of customer satisfaction on previous or ongoing contract efforts, number of video conference contacts, number of in-person contacts and the number of labor hours devoted to key systems engineering activities such as requirements, architecture and cost estimation. The factor relationship study also includes a set of recommendations for people using systems engineering on proposals to improve their chances of being awarded contracts.

This factor analysis answers the call for more empirical research in systems engineering [Valerdi and Davidz, 2008]. This factor analysis also contributes to the proposal management literature. No study was found where factors of any kind (systems engineering related or not systems engineering related) were quantitatively examined in relationship to the probability of a supplier being awarded a contract. Because the focus of the factor analysis study presented in this dissertation is systems engineering related, there may be other factors that are important for a more general proposal management study that were not considered in this research.

5.3.2 Validated Optimization Modeling Framework for Systems Engineering on Proposals

Another important contribution is the optimization modeling framework (i.e., SEPOMF). SEPOMF is designed for the use of systems engineering on proposals. The SEPOMF can be used by organizations where organizations develop models with their historical data. The SEPOMF provides explicit directions and guidance for organizations to optimize how they use systems engineering on proposals. For a complete description of what the SEPOMF is and how to use the SEPOMF, see Appendix G.

5.4 Future Work

The research published in this dissertation plus the papers by Smartt and Ferreira [2011, 2012c] and Philbin [2008] are the beginning of research related to the use of systems engineering on proposals. Because organizations that thrive on contract work need to be successful at the proposal process to survive, because systems engineering plays such a central role in proposal management and because there is potential competitive advantage to be gained by judiciously using systems engineering on proposals, more research should continue in this area. Some of this research should be performed in scholarly settings where researchers generally have significant flexibility and freedom, while other research should be performed within organizations where researchers sometimes have access to repositories rich with project data. Hopefully, organizations will attempt to apply the concepts from the scholarly studies such as this one using their internal data. Ideally, there could be some feedback from these organizations to the academic community regarding the level of success organizations enjoy from such approaches. In this section, the discussion of future work is divided into several categories. These include: (1) the next steps to continue the research trajectory initiated through this dissertation, (2) more precisely focused future research relating to the use of systems engineering in proposal management, (3) more broadly focused research relating to the use of systems engineering in proposal management, and (4) future research in areas tangential to use of systems engineering in proposal management.

5.4.1 Next Steps in Current Research Trajectory

The next future step in this research should be collecting a sufficiently large set of historical proposal efforts to actually develop a data-driven DSS using the SEPOMF. If this is done successfully (i.e., the objective function makes sense, the optimal solutions found by the model are reasonable), then the level of confidence in the utility of this concept will increase. In addition, the very technical explanations for "how to" use the SEPOMF in this dissertation can be potentially simplified and expanded to be usable by people with less depth in analytical modeling. This simplification may come in the form of written directions, automated tools for model creation or a combination of the two. Future research in this area, however, should not be limited to simply maturing the concept presented in this dissertation.

5.4.2 More Precisely Focused Future Research

Some of this future research may be more precisely focused and some of the research may be broader in scope. The more focused research may go more in depth related to particular factors that were found to be statistically correlated with an increased probability of contract award. This future research may attempt to gather larger quantities of data so that the statistical relationships can be better understood. The future research may also consider factors that were not explicitly considered in this dissertation but subsequent feedback from multiple subject matter experts has indicated to be of potential significance, such as level of risk. The more focused research may also address how systems engineering can be used to improve performance related to each factor in a way that will result in attaining more contract awards. For example, it would be helpful to understand how to use systems engineering to satisfy existing customers.

5.4.3 More Broadly Focused Future Research

Other research may take a broader perspective and examine the role of systems engineering in proposal management more from the perspective of long term sustainment of an organization. Research related to using systems engineering to optimize other proposal objectives [Smartt and Ferreira, 2012c], such as net present value of a proposal opportunity, the value of a real option related to the proposal opportunity and the capital position of an organization may prove very valuable in helping organizations leverage systems engineering to attain higher-level goals.

When examining broader objectives, the research should seek to optimize a portfolio of projects versus just an individual proposal project. Organizations generally manage a portfolio of different projects. Organizations have limited resources and a limited number of available labor hours for professionals with certain skill sets. In general, allocating those resources and assigning those individuals to one project in the portfolio by definition makes them unavailable for other projects in the portfolio. Those other projects could be projects to execute an already captured contract, projects to enhance the organization’s capabilities, or projects to pursue other contracts. Allocating resources and assigning personnel to a proposal effort of interest will make those resources and individuals’ time unavailable for ongoing contract work, unavailable to enhance the organization’s capabilities and unavailable for proposal efforts other than the proposal of interest.

5.4.4 Tangentially Focused Research

Some of the work performed on this dissertation provides opportunities for future research beyond the research area of the use of systems engineering on proposals. For example, a number of factors with potential importance to the use of systems engineering on proposals that have been identified in this research may have broader applications to systems engineering strategy. A useful research contribution would be to use a relevant subset of the factors identified in this research to enhance the systems engineering strategy framework described in Smartt and Ferreira [2012a] and refined in Smartt and Ferreira [2012b]. Another future research project could be to mature the systems engineering skill level rubric used in this research, reconcile it with other models [Davidz and Nightingale, 2008; INCOSE, 2010] and potentially validate the rubric. The existence of such a rubric may help professional in systems engineering plan their career paths so that they can more systematically advance their careers. Also, the methodology used to validate optimization modeling frameworks should be refined and validated. A validated methodology for an optimization modeling framework may prove useful for future research seeking to define an optimization model but lacking the actual data to apply the model at the time of model definition.

Appendix A

Survey

Definitions and Directions

Survey Definitions

The following definitions are provided to the survey respondents in order to ensure a consistent meaning for the key terms used in this survey.

Systems engineering, as defined in this survey, is: "an interdisciplinary approach and means to enable the realization of successful systems" [INCOSE, 2004; p. 12]. It focuses on defining customer needs and required functionality early in the development cycle, documenting requirements, and then proceeding with design synthesis and system validation while considering the complete problem: operations, cost and schedule, performance, training and support, test, manufacturing, and disposal. SE considers both the business and the technical needs of all customers with the goal of providing a quality product that meets the user needs.

Development, as defined in this survey, includes all of the effort from when an organization is awarded a contract until the system is delivered to a customer. This includes all steps from concept definition through transition to use.

Survey Directions

Field tests of the survey reveal that it takes approximately 30 minutes to complete, excluding any time searching for data in project documents. The survey contains 38 questions.

Some of the questions ask for specific quantities of interest. It is acceptable if you do not have exact answers for many of these questions. This survey seeks to capture your best estimate. For the purposes of this research, your best estimate is much better than no answer at all.

If you have knowledge of past proposal efforts that meet the following criteria, please answer this questionnaire using information from your most recently submitted proposal for which an award status is known. If you have knowledge of more than one qualifying proposal effort, we welcome multiple submissions, one submission for each proposal effort. Types of proposal efforts of interest generally relate to those pursuing system development contracts that deliver new desired functionality. These types of contracts include:

1. Introduction of a new system for which there is no predecessor (e.g., the first generation of cell phones, the first generation of microwave ovens)

2. Replacement of an old system in which the new system provides new desired functionality that was not provided by the old system (e.g., replacing a traditional cell phone with a smart phone)

3. Enhancement to an existing system that results in new desired functionality

The following criteria must all be satisfied to enter data related to a proposal:

1. You must have knowledge of how systems engineering was used on a proposal effort. As the focus of this research is on the systems engineering function, please respond regardless of whether or not the titles of the individuals performing the tasks include “systems engineer”.

2. **The contract award status must be known**. In other words, you must know definitively if a contract was awarded or not as a result of the proposal effort. Proposal efforts resulting in contract awards and those where no contract was awarded are both of interest. Contracts where it remains to be determined whether an award will be granted are not of interest.

3. The proposal must be for a contract that includes development (as defined in the survey definitions). Contracts of interest include any system that will ultimately be delivered to a customer, whether a one of a kind system or a mass produced system. Contracts to support operations, maintenance and disposal of systems are not the focus of this survey. If a contract includes both development and support for the system after development, the development portion of the contract is of interest. Contracts to deliver systems that are already defined and contracts for research efforts where the end deliverable is a report are not of interest.

4. The proposal must have been submitted on or after January 1, 2007.

The following is the survey’s agenda and lists the major question categories:

1. Your background, experience and role in the proposal

2. Information about the contract being pursued

3. Information about the proposal effort

Validation and Contact Information

1. In order to validate responses and capture contact information for the gift certificates offered, you are asked to provide your name, either a contact email address or phone number and membership status related to particular professional organizations. You will not be contacted unless you win a prize or request to be contacted. All individual survey results will be kept confidential and not associated with an individual or organization.

- Name
- Email
- Phone Number

2. **I am a member of or the organization that employs me is a member of (select all that apply):**

- International Council on Systems Engineering (INCOSE)
- National Defense Industrial Association (NDIA)
- Project Management Institute (PMI)
- None of the above

3. **Unique Project Identifier -** As stated above, you are welcome to enter data for multiple qualifying proposal efforts. This form should be filled out once for each unique proposal effort. Because of this, you are asked to provide a unique identifier for the project.

Outcome of Proposal Effort

4. **The outcome of the proposal effort was (select one):**

a. A contract was awarded to my organization.
b. No contract was awarded to my organization, but a contract was awarded to another organization.
c. No contract was awarded to any organization and no contract is expected to be awarded.
d. Unknown

Your Experience and Background

5. **How many total years of work experience do you have?** Please provide the full-time equivalent of the number of years that you have worked. Please count the experience for all of the different employers you have worked for. For example, if you have worked half-time for ten years, the answer should be 5 years.

6. Your experience in relation to a number of different roles is assessed in this section. It is not expected that you served in all of these roles. **For the roles below, please provide the equivalent number of years that you have served in each.** For example, if you worked full-time for 5 years, 60% of the time working as an individual contributor with systems engineering as your primary responsibility and 40% of the time working as a chief systems engineer, then you have worked 3 years as an individual contributor with systems engineering as your primary responsibility and 2 years as a chief systems engineer as your primary responsibility. Part of a year can be represented by a decimal. The total for your answers to the following should sum to the answer of the previous question.

- Number of years (full-time equivalent) served as an individual contributor with systems engineering as my primary responsibility

- Number of years (full-time equivalent) served as an individual contributor in an engineering discipline other than systems engineering as my primary responsibility (e.g., electrical engineer, software engineer)
- Number of years (full-time equivalent) served in the role of chief or lead systems engineer as my primary responsibility
- Number of years (full-time equivalent) served in the role of chief engineer or lead engineer as my primary responsibility
- Number of years (full-time equivalent) served in the role of project manager as my primary responsibility. This includes managing proposal projects as well as managing projects that execute contracts that have already been awarded
- Number of years (full-time equivalent) served in the role of program manager as my primary responsibility. A program manager oversees a portfolio of projects.
- Number of years (full-time equivalent) served as a functional manager responsible for a group of individuals performing systems engineering as my primary responsibility
- Number of years (full-time equivalent) served as a specialist in capturing business (e.g., capture manager)
- Number of years (full-time equivalent) served as a director (e.g., Director of Engineering, Director of Marketing, Director of Program Management)
- Number of years (full-time equivalent) served as an executive officer or vice president
- Number of years (full-time equivalent) served in roles other than the ones listed above

7. **If you served in roles other than one of those described above, please list each role, years in each role (full-time equivalent) and a brief description of each role:**

Information About the Contract Being Pursued

8. **The customer organization's home office is located (select one):**

- In the United States
- Outside the United States

9. **The end user for the system is from the _______ sector (select all that apply).**

- Commercial
- Military/defense
- Government/non-military

10. **The type of system can be categorized as ____ (select all that apply).**

- Banking/financial system (e.g., ATM, database of financial records)

- Communication system (e.g., landline infrastructure, baby monitor)
- Electronics system (e.g., personal communication devices, gaming systems)
- Healthcare system (e.g., MRI machine, automated intravenous pharmaceutical dispensing device)
- Information system (e.g., network of computers providing users information by accessing data stores)
- Infrastructure system (e.g., bridges, dams)
- Shipping/distribution/logistics system (e.g., warehouse management)
- Training system (e.g., simulators)
- Transportation system (e.g., automobile, urban mass transit system, ship, aircraft)
- Weapon system (e.g., missile fire control, military intelligence system, fighter aircraft, central command operations center)
- Other type of system

11. **If other type of system, please describe:**

12. **The type of system required to satisfy the intended contract could best be described as (select one):**

a. A new system that is the first of its kind
b. A new system to replace an existing system (but provides new functionality)
c. An enhancement or upgrade to an existing system (that provides new functionality)

13. **The type of contractual arrangement that was applicable or would have been applicable if your organization had been awarded a contract could best be described as (select one):**

a. Firm, fixed price - provides for a price that is not subject to any adjustment on the basis of the contractor's cost experience in performing the contract [US, 2013].
b. Cost plus fixed fee - is a cost-reimbursement contract that provides for payment to the contractor of a negotiated fee that is fixed at the inception of the contract. The fixed fee does not vary with actual cost, but may be adjusted as a result of changes in the work to be performed under the contract [US, 2013].
c. Cost plus percentage of cost - is a cost-reimbursement contract in which the contractor receives the costs expended plus a fee proportional to the overall costs.
d. Other type of contracting arrangement

14. **If other type of contracting arrangement, please describe:**

15. **The organization submitted a proposal in order to (select all that apply):**

- Make a profit from executing the development portion (as defined in the survey definitions) of the contract
- Make a profit from operating, maintaining or disposing of the system after delivery
- Increase the probability of being considered for future opportunities involving the same customer

- Establish a business relationship with other supplier organizations
- Satisfy an existing customer who expected a response
- Deny a potential competitor an opportunity

16. **When your organization submitted the proposal, how many units of the system (e.g., quantity of systems) was the organization planning to deliver to the customer?**

The following questions pertain to the amount of development work that would be performed on contract if a contract was to be awarded. Recall that development, as defined in this survey, includes the portion of the system lifecycle from when an organization is awarded a contract until the system is delivered to the customer. Please answer these questions regardless of whether your organization was awarded a contract or not.

17. **When the decision was made to prepare a proposal, it was estimated that the potential contract would provide a total of ______ systems engineering (as defined in the survey definitions) labor hours of work related to development (as defined in the survey definitions) to my organization.** For example, if the development portion of the contract was estimated to support 3 people performing systems engineering for 5 years at 1,700 labor hours per person each year, the answer would be 5 years*3 people performing systems engineering/year*1,700 labor hours/person performing systems engineering = 25,500 labor hours. If some of the systems engineering effort in the contract pertained to supporting the system after delivery to the customer, please do not include that portion of the contract.

18. **When the decision was made to prepare a proposal, it was estimated that the potential contract would provide a total of ______ engineering labor hours (all engineering including systems engineering) of work related to development (as defined in the survey definitions) to my organization.** For example, if the contract was estimated to support 30 engineers for 5 years doing development work at 1,700 hours per engineer for each year, the answer would be 5 years*30 engineers/year*1,700 labor hours/engineer = 255,000 labor hours. If some of the engineering effort in the contract pertained to supporting the system after delivery to the customer, please do not include that portion of the contract.

19. **The development portion of the contract being pursued was estimated to take _____ months of calendar time.** This equates to duration. In other words, if the development effort was 2 years, the answer is 24 months. Duration is different than total labor effort.

Information about Proposal Effort

The following questions pertain to the proposal effort in pursuit of the contract.

20. **What calendar year (e.g., 2009) was the proposal submitted in?**

21. **Your role in the described proposal effort could be described as (select all that apply):**

- Systems engineering individual contributor – I did systems engineering work on the proposal effort
- Engineering individual contributor for a discipline other than systems engineering – I did engineering work on the proposal effort, but in a discipline other than systems engineering
- Chief or lead systems engineer on proposal effort
- Chief or lead engineer
- Project manager of proposal effort
- Program manager of a portfolio of projects including the proposal effort
- Functional manager responsible for a group of individuals performing systems engineering
- Specialist in capturing business (e.g., capture manager)
- Director (e.g., Director of Engineering, Director of Marketing, Director of Program Management) Executive officer or vice president
- Other role(s)

22. **If you served in another role(s), please list the role(s) and a brief description of each role:**

One key set of activities in a proposal effort that systems engineering often leads is defining the proposed offering. Defining the proposed offering relates to defining the requirements and technical solution that will be offered in the proposal. This survey focuses on the following core activities: define and validate source requirements, define and validate system technical requirements, define and validate system architecture alternatives, estimate costs for each alternative and perform decision analysis to select a final solution.

There are some systems engineering related activities that may occur during a proposal effort that are not included in the core activities for defining the proposed offering. For example, the process of explicitly planning the execution of the project (including defining a work breakdown structure, defining a master schedule, defining a systems engineering management plan) is not included because it is not in the scope of this particular research analysis.

The information below provides a detailed definition of the core activities in defining the proposed offering.

Define and validate source requirements - The source requirements are all requirements and constraints imposed by the customer and other stakeholders. This includes defining the source requirements and validating them with the customer or other stakeholders.

The following processes are included: acquirer and other stakeholder requirements definition, acquirer and other stakeholder requirements validation, acquirer and other stakeholder requirements negotiation and analysis, management of acquirer and other stakeholder requirements, time spent following an already defined configuration management process in support of managing source requirements, decision analysis (e.g., trade studies) related to acquirer or other stakeholder requirements and identifying and evaluating risks related to acquirer and other stakeholder requirements.

Define and validate system technical requirements - This activity involves analyzing the source requirements (which are often stated in user or operational terms) and formulating a set of system-level technical requirements. This also includes validating the system technical requirements with the customer or other stakeholders.

The following processes are included: system technical requirements definition, system technical requirements validation, system technical requirements negotiation and analysis, creating traceability between system technical requirements and source requirements, management of system technical requirements, time spent following an already defined configuration management process in support of system technical requirements, decision analysis (e.g., trade studies) related to system technical requirements, and identifying and evaluating risks related to system technical requirements.

Define and validate architecture alternatives - This involves defining alternative system architectures at both a logical and physical level. This also includes identifying and analyzing key technologies for the architecture and ensuring that requirements for the identified system are addressed by each architecture.

The following processes are included: architectural design process, definition of both logical and physical architectures, high-level system design (process of defining the major elements in the design and the arrangement of the elements), assessment of technologies that may be inserted into the system, validation of both logical and physical architectures, time spent following an already defined configuration management process in support of architecture, decision analysis (e.g., trade studies) related to architecture and technology insertion, and identifying and evaluating risks related to system architecture.

Estimate cost for each alternative - This includes using various approaches to estimate the relevant costs for each alternative.

The following processes are included: time spent “sizing” the system in terms of attributes that affect the labor hour costs and material costs for delivering a system. Labor hour cost drivers may include number of requirements, number of interfaces, number of critical algorithms and number of lines of source code or number of function or feature points for software intensive systems. Material cost drivers may include hardware or software packages (e.g., COTS) that must be acquired in order to build the system. Estimating the cost for each alternative also includes time spent using tools and processes to generate cost estimates for various solution alternatives, time spent following an already defined configuration management process in support of cost estimation, and identifying and evaluating risks associated with cost estimates.

Perform decision analysis to select final solution - This may involve applying decision analysis (e.g., trade study) to evaluate the various solution concepts and selecting a preferred concept. This includes time spent defining criteria, weights, evaluating alternatives against criteria, sensitivity analysis and selecting one of multiple defined solution concepts.

23. **How many different people in your organization were involved in systems engineering activities related to the core activities (see table above) when defining the proposed offering?** Please only include people who directly contributed, and do not include people who served on the proposal exclusively in a review capacity.

For each core activity undertaken when defining the proposed offering for a proposal effort, employees of different skill levels are assigned. The various skill levels in this survey are defined by the table below. Levels are defined with respect to formal education level, number of years of domain experience and number of years performing a particular activity. Please consider all three when determining an employee's skill level. The number of years of domain experience is defined as the number of years that the employee has been working with similar systems or similar technologies. Time spent in roles such as project management or engineering (including systems engineering) all count toward domain experience.

Skill Level	Formal Education Level	Number of Years of Domain	Number of Years of Experience
Expert	No college degree	20+ years	6+ years
	Associate degree or bachelor's degree in non-technical field	16+ years	6+ years
	Bachelor's degree in technical field	12+ years	6+ years
	Masters or PhD in technical field	8+ years	6+ years
Advanced	No college degree	16+ years	6+ years
	Associate degree or bachelor's degree in non-technical field	12+ years	6+ years
	Bachelor's degree in technical field	8+ years	6+ years
	Masters or PhD in technical field	4+ years	4+ years
Intermediate	No college degree	12+ years	6+ years
	Associate degree or bachelor's degree in non-technical field	8+ years	6+ years

	Bachelor's degree in technical field	4+ years	4+ years
Beginner	Does not meet requirements for intermediate		

24. **Identify how many labor hours were allocated to each activity and skill level for defining the proposed offering for the proposal effort segmented into activities.** Using the above definitions for skill level and core activities, please fill out the following matrix. All entries should be in labor hours. If a person was assigned full-time to a particular activity for 2 months, then the value would be approximately equal to 8 hours/day*5 days/week*4 weeks/month*2 months = 320 hours for the corresponding entry. Note that in many cases, entries may have values of zero. For example, if only an advanced employee was entrusted with all source requirement definition and validation, then the entries for beginner, intermediate and expert level people for source requirements should contain zeros. All answers reflect what occurred in your organization only.

- Define and validate source requirements - Hours allocated to employees with a Beginner skill level
- Define and validate source requirements - Hours allocated to employees with an Intermediate skill level
- Define and validate source requirements - Hours allocated to employees with an Advanced skill level
- Define and validate source requirements - Hours allocated to employees with an Expert skill level
- Define and validate system technical requirements - Hours allocated to employees with a Beginner skill level
- Define and validate system technical requirements - Hours allocated to employees with an Intermediate skill level
- Define and validate system technical requirements - Hours allocated to employees with an Advanced skill level
- Define and validate system technical requirements - Hours allocated to employees with an Expert skill level
- Define and validate architecture alternatives - Hours allocated to employees with a Beginner skill level
- Define and validate architecture alternatives - Hours allocated to employees with an Intermediate skill level
- Define and validate architecture alternatives - Hours allocated to employees with an Advanced skill level
- Define and validate architecture alternatives - Hours allocated to employees with an Expert skill level
- Estimate cost for each alternative - Hours allocated to employees with a Beginner skill level
- Estimate cost for each alternative - Hours allocated to employees with an Intermediate skill level
- Estimate cost for each alternative - Hours allocated to employees with an Advanced skill level

- Estimate cost for each alternative - Hours allocated to employees with an Expert skill level
- Perform decision analysis to select final solution - Hours allocated to employees with a Beginner skill level
- Perform decision analysis to select final solution - Hours allocated to employees with an Intermediate skill level
- Perform decision analysis to select final solution - Hours allocated to employees with an Advanced skill level
- Perform decision analysis to select final solution - Hours allocated to employees with an Expert skill level

The following questions relate to process maturity. Process maturity in this survey is rated in accordance with how the Software Engineering Institute (SEI) has defined maturity levels for the staged representation in Capability Maturity Model Integration (CMMI) for Development v1.3 dated November, 2010 [SEI, 2010]. URL: http://www.sei.cmu.edu/reports/10tr033.pdf

These are summarized as follows:

CMMI Level	Name	Description
1	Initial	The processes within the organization are ad-hoc and chaotic with ineffective management procedures and project plans.
2	Managed	The processes are planned and executed in accordance with policy; the projects employ skilled people who have adequate resources to produce controlled outputs; involve relevant stakeholders; are monitored, controlled, and reviewed; and are evaluated for adherence to their process descriptions.
3	Defined	The organization uses defined processes on all of its projects. A defined process is a managed process that is tailored from the organization's set of standard processes according to the organization's tailoring guidelines.
4	Quantitatively Managed	The organization and projects establish quantitative objectives for quality and process performance and use them as criteria in managing projects.

5	Optimizing	The organization continually improves its processes based on a quantitative understanding of its business objectives and performance needs. The organization uses a quantitative approach to understand the variation inherent in the process and the causes of process outcomes.

25. **What is the process maturity of your organization (select one)?** If your organization has a CMMI rating, please use that rating. If your organization has not received a CMMI rating, provide your best judgment based upon the above definitions.

a. CMMI Level 1: Initial
b. CMMI Level 2: Managed
c. CMMI Level 3: Defined
d. CMMI Level 4: Quantitatively Managed
e. CMMI Level 5: Optimizing

26. CMMI is defined at the organization level. **However, if I was to assign a CMMI rating at the project level for the proposal, it would be (select one).** The answer may be the same or different than the previous question.

a. CMMI Level 1: Initial
b. CMMI Level 2: Managed
c. CMMI Level 3: Defined
d. CMMI Level 4: Quantitatively Managed
e. CMMI Level 5: Optimizing

Information About Prime Contractor

It is common for teams of contractors to collaborate on a proposal effort with the expectation that if a contract is awarded, all collaborators will win business. In such arrangements, there is usually a prime contractor who contracts directly with the customer, and the other contractors are sub-contractors to that prime.

The following questions pertain to the prime contractor for the proposal effort, whether that was your organization or not. If your organization was a subcontractor, but you have answers or estimates to these questions, please input your information.

27. **In the proposal effort, my organization was (select one)**:

a. The prime contractor

b. A subcontractor to the prime contractor

Information About Prime Contractor

28. **What is the process maturity of the prime contractor organization (select one)?** If the prime organization has a CMMI rating, please use that rating. If the prime organization has not received a CMMI rating, provide your best judgment based upon the previously provided definitions.

a. CMMI Level 1: Initial
b. CMMI Level 2: Managed
c. CMMI Level 3: Defined
d. CMMI Level 4: Quantitatively Managed
e. CMMI Level 5: Optimizing

29. **During the proposal stage (time from when the prime contractor decided to pursue a contract opportunity until the proposal was submitted), the prime contractor had (select one):**

a. No access to the customer
b. Authorization to interact with the customer a set number of times as specified by the customer
c. Authorization to interact with the customer whenever the prime felt the need to interact
d. Other type of customer interaction

30. **If you had another type of customer interaction (including a combination of previously described types of interactions), please describe:**

The next question pertains to the number of interactions between the prime contractor organization and the organization acting as acquirer from the time when the prime contractor decided to pursue a contract opportunity until the proposal was submitted. Some interactions between suppliers and acquirers involve multiple forms of communication used simultaneously. Please count each interaction only once and count it as the highest type of interaction according to the matrix below. Interaction types are listed in ascending order. For example, if there was a real-time, face-to-face meeting between someone from the acquirer’s organization and someone from the supplier’s organization where the two parties are co-located (i.e., Level 4 Interaction) and other people from both organizations were teleconferenced in (Level 3 Interaction), please count the meeting as one Level 4 interaction. Do not also count it as a Level 3 Interaction.

Level	Name	Description
Level 1	Written	Written exchanges related to the proposal effort (e.g., email, traditional mail, texts)
Level 2	Real Time, Voice-Only, Remote	Interactions related to the proposal effort via real-time, voice only communications (e.g., telephone calls, cell phone calls)
Level 3	Real-Time, Remote	Interactions related to the proposal effort using technology that enables participants to see and hear each other (e.g., video conferencing, using an application like Face Time or Skype)
Level 4	Real-Time Co-located	In-person interactions related to the proposal effort where representatives from the supplier and acquirer interact in the same room

31. **During the proposal stage, the prime contractor and the customer had:**

- Number of Level 1 Interactions
- Number of Level 2 Interactions
- Number of Level 3 Interactions
- Number of Level 4 Interactions

32. **The overall leader (project manager) of the proposal effort for the prime contractor had (select one):**

a. Led more than one proposal effort prior to the proposal effort described here
b. Led exactly one proposal effort prior to the proposal effort described here
c. Never been an overall lead (project manager) on a proposal effort prior to the proposal effort described here, but was highly experienced in proposal efforts as either a contributor or a leader other than the overall leader (e.g., chief systems engineer, IPT lead)
d. Never been an overall lead (project manager) on a proposal effort prior to the proposal effort described here, but had some experience in proposal efforts as either a contributor or a leader other than the overall leader (e.g., chief systems engineer, IPT lead)
e. Had no experience with proposal efforts prior to the proposal effort described here
f. Unknown

33. **The competitive rank of the prime contractor in revenue from sales could be described as __ (select one).** Competitive rank in terms of revenue from sales relates to the percentage of revenue the organization earns for comparable products in the entire marketplace, including those made by competitors. If, for example, the organization has greater revenue from sales than any of its competitors, then it ranks in the 75-100

percentile. Note that this is not market share. If there are many competitors, an organization may earn a small percentage of the total revenue, but still be highly ranked if other organizations have even less market share.

a. 75 – 100 percentile
b. 50 – 74 percentile
c. 25 – 49 percentile
d. 0 - 24 percentile
e. Unknown

34. **The competitive rank of the prime contractor in prestige could be described as (select one).** Competitive rank in terms of prestige is defined as how an organization would rank in terms of customer preferences if money was not an object. For consumer goods, luxury products would be examples of products that enjoy a high competitive rank in terms of prestige. For example, if most drivers would prefer to drive a luxury automobile instead of other types of vehicles, money not being an object, then the luxury automobile would rank in the 75 to 100 percentile with respect to prestige.

a. 75 – 100 percentile
b. 50 – 74 percentile
c. 25 – 49 percentile
d. 0 - 24 percentile
e. Unknown

35. **What is the extent of the previous relationship (including both contracts and contacts) between the prime contractor and the customer?** The prime contractor (select one):

a. Was executing a separate contract with the intended customer at the time the proposal effort was submitted
b. Had executed a separate contract with the intended customer within a two year period prior to the submission of the proposal, but was not executing a contract with the customer at the time of the proposal submission
c. Had no separate contracts with the customer within a two year period, but did have past contacts with the customer
d. Has had no existing previous contacts or contracts with the intended customer

Information About Prime Contractor

36. **In past or ongoing contracts with the prime contractor at the time of the proposal submission, this customer's general level of satisfaction with the performance of the prime contractor could be best described as (select one):**

a. Very satisfied
b. Somewhat satisfied
c. Neither satisfied nor dissatisfied
d. Somewhat dissatisfied

e. Very dissatisfied

Confidence in Your Selections

37. **What is your level of confidence in the qualitative selections you supplied for this survey (select one)?**

a. Extremely confident
b. Very confident
c. Confident
d. Somewhat confident
e. Not confident at all

38. **What is your level of confidence in the quantitative (numeric data) inputs you supplied for this survey (select one)?**

a. Extremely confident
b. Very confident
c. Confident
d. Somewhat confident
e. Not confident at all

Appendix B

Survey Verification

B.1 Objectives

This appendix provides an overview of the testing conducted before the Survey Monkey version of the *Survey to Assess the Relationship between Systems Engineering Factors and Proposal Success* went live. A similar plan was defined and executed before the Microsoft Expressions version of the survey went live. The verification plan for the Survey Monkey version of the survey is included because it is slightly more refined than the plan for the Microsoft Expressions version of the survey.

After executing this verification plan and attaining the desired results:

1. There should be high confidence that the capability exists to support the collection of the data needed for the survey.
2. There should be high confidence that technical issues will not occur that will keep this data from being properly stored and available for analysis.
3. There should be high confidence that unauthorized people will not have access to this data.

The primary focus of this test activity is verification of the test instrument and related infrastructure such as operating system and browser software.

B.2 Scope

The scope of this appendix describes the complete testing plan that is needed to comprehensively test the survey instrument. This includes:

1. Testing the actual electronic survey forms (e.g., correctness, look/feel)
2. Testing the capability and process for recording and storing user input data
3. Testing to ensure surprise behavior does not occur
4. Testing to see that form and function work properly across multiple browsers and operating systems

B.3 What is Being Tested

The unit under test is a survey that is coded using the tools and services of Survey Monkey. Survey Monkey is a service that supports the creation, hosting and analysis of results for surveys. Users who fill out this survey will have their data recorded in a database. The results from that database are downloaded by the researcher and analyzed in conjunction with data collected from the Microsoft Expressions version of the survey originally deployed.

B.4 Verification Approach

This section describes the types of tests that will occur, the documentation that will be created and the plan for regression testing in the event that modifications to the survey are made in the review process.

B.4.1 Types of Tests

There were a number of different types of tests conducted, including tests created for the web site creator (i.e., the researcher) to conduct, tests by dissertation committee members, tests by peer reviewers and tests by experts in the field. While the expert field tests were primarily conducted for the purpose of validating the survey content, they also served as an extra layer of verification to ensure that the survey was error-free.

B.4.1.1 Web Site Creator Conducted Tests

These are tests conducted by the website creator.

B.4.1.1.1 Form Unit Tests

1. Verify radio buttons work as expected
2. Upon opening the form, no radio buttons are selected by default.

3. Upon clicking any radio button, that radio button becomes selected and any other radio button in that group that was previously selected becomes deselected.
4. Selecting any particular radio button results in a unique value being recorded to the database that maps to that particular selection upon progressing to the next page of the survey.
5. Verify text boxes and text areas work as expected.
6. Users may input text into any text field.
7. User may erase text from any field.
8. All text in text fields is recorded to the database upon progressing to the next page of the survey.
9. Establish that the number of allowable characters that may be entered in any of these fields is adequate.
10. Verify that at least this number of characters is recorded to the output file upon progressing to the next page of the survey.
11. Verify check boxes work as expected.
12. Upon opening the form, no checkboxes are checked.
13. The user may elect to check any box.
14. The user may elect to uncheck any box.
15. For any checkbox, whether the box is checked or not is recorded to the database upon progressing to the next page of the survey.
16. Verify hyperlinks work.
17. Verify hyperlinks in all emails or website posts announcing the survey actually direct the user to the survey.

18. Verify that all hyperlinks in the survey to web sites outside the survey take the user to the intended locations.
19. Test question logic.
20. Ensure that the respondent is sent along the correct paths given their answers to key questions that invoke question logic.

B.4.1.1.2 Form Content Reviews

1. Verify spelling is correct.
2. Verify formatting is correct. All information is aligned in a consistent and aesthetically pleasing way.
3. Verify that all explanatory information in the Microsoft Expressions survey is included in the Survey Monkey version of the survey.
4. Verify that all questions in the Microsoft Expressions survey are also included in the Survey Monkey version of the survey.

B.4.1.2 Performing Tests in Different Environments

There are now a very large number of operating systems, a large number of browsers, and a very large number of potential combinations of operating systems and browsers. Comprehensive testing is impractical. What will be done instead is testing of key combinations of operating systems and browsers. All of the Web Site Creator Conducted Tests are to be performed in the following combinations of browsers and operating systems:

1. Windows Explorer in Windows Premium Home OS: Windows 7 Home Premium/Internet Explorer 9.0.8112.16421
2. Firefox in Windows Premium Home OS: Windows 7 Home Premium/Firefox 13.0

3. Chrome and Windows Premium Home OS: Windows 7 Home Premium/Chrome 19.0.1084
4. Windows Explorer on a Windows VISTA OS: Windows VISTA Premium Home/Internet Explorer 9.0.8112.16421
5. Safari in Mac OS on a personal computer: Mac OS 10.5.8/Safari 5.0.6
6. Safari on iPad 2: iPad2 (running Mac CPU OS 4_3_5)/Safari 6533.18.5
7. Firefox in Linux OS: Linux 3.4.3-1.fc17/Firefox 13.0.1

B.4.1.3 Committee Member Tests

Committee members were informed once the decision was made to reformat to Survey Monkey. They were asked to indicate if they wished to test the reformatted survey. Only the dissertation supervisor tested the Survey Monkey version of the survey.

B.4.1.4 Peer Tests

These are tests to be conducted by lay people with varying degrees of expertise related to computer systems and the use of the internet. Two people participated in peer testing the Survey Monkey version of the survey. Previously, four peers had tested the Microsoft Expressions version of the survey.

B.4.1.5 Field Tests by Subject Matter Experts

Five individuals with appropriate credentials participated in testing the Microsoft Expressions version of the survey. As the focus on their review was primarily on content, there was not really a need to repeat field tests for the Survey Monkey reformat of the survey. Instead, as previously mentioned, care was taken to ensure all directions and questions from the field-tested Microsoft Expressions version of the survey were translated intact to the Survey Monkey version of the survey.

B.4.2 Test Artifacts

Test artifacts include documentation of form unit tests, documentation of form content reviews, documentation of peer tests and committee member tests and a regression testing plan. A matrix was filled out where each test was featured on a row and each operating system/browser combination was featured in a column. A check signifies that the criteria in row *i* has been successfully demonstrated in browser/operating system combination in column *j*. A blank version of this test sheet is featured in Table B.1.

Form content reviews were conducted only one time as the content is operating system/browser agnostic. The form content was documented by a populated check list. Regression testing refers to any testing that has to be repeated because of changes that are made to the survey as a result of findings from other testing.

All web site creator conducted tests had to be successfully passed before peers were asked to test.

Depending upon the significance of any changes resulting from the peer tests, some or all of the web site creator conducted tests were repeated on the updated webpage.

Table B.1 Survey Verification Matrix

	Form Unit Tests	Win 7 Home/ IE	Win 7 Home/ FF	Win 7 Home/ Chr	Win VISTA /IE	Mac Pcor /Saf	Mac iPad/ Saf	Linux/ FF
2	Verify radio buttons work as expected							
3	1.) Upon opening the form, no radio buttons are selected by default.							
4	2.) Upon clicking any radio button, that radio button shall become selected and any other radio button in that group is selected is deselected.							
5	3.) Selecting any particular radio button results in a unique value being recorded to the database that maps to that particular selection upon hitting SUBMIT.							
6	Verify text boxes and text areas work as expected							
7	1.) Users may input text into any text field.							
8	2.) User may erase text from any field.							
9	3.) Text in text fields are recorded to the database upon the user hitting SUBMIT.							
10	Verify check boxes work as expected							
11	1.) Upon opening the form, no checkboxes are checked.							
12	2.) The user may elect to check any box.							
13	3.) The user may elect to uncheck any box.							
14	4.) For any checkbox, whether the box is checked or not is recorded to the database upon the user hitting submit.							
15	Test that "DONE" button							
16	1.) Upon selecting DONE, the user is presented with a thank you page.							
17	2.) All values in the form are recorded/appended to a database and encoded as intended.							
18	Verify Hyperlinks							
19	1.) Verify link to survey works							
20	2.) Verify link to Adobe works							
21	3.) Verify link to INCOSE works							
22	4) Verify link to DAF works							
23	5.) Verify link to CMMI works							
24	*Form Content Reviews*							
25	1.) Verify spelling is correct.							
26	2.) Verify formatting is correct. All information is aligned in a consistent and aesthetically pleasing way.							
27	3.) Verify that all explanatory information in .docx is covered. In other words, ensure that no content in the developmental .docx version of the survey has been accidentally omitted.							
28	4.) Verify that all questions in .docx version of survey are covered. In other words, ensure that no content in the developmental .docx version of the survey has been accidentally omitted.							
29								

Appendix C

Survey Validation

C.1 Validity Overview

The *Survey to Assess the Relationships Between Systems Engineering Factors and Proposal Success* has the following objectives: (1) Obtain a description of how organizations are using systems engineering in proposal management, (2) Determine which of the systems engineering related factors related to proposal management are statistically significant, and (3) Attempt to provide the statistical data to support the DSS. In order to achieve these goals, it is important that the survey have a high level of validity. Surveys are valid that measure what they intend to measure [Litwin, 1995].

There are different types of validity to consider when designing a survey. Two main categories of types of validity for a survey are internal and external. Findings from a study have internal validity if effects observed in the dependent variable are actually caused by the independent variable and not by other factors [Rubin and Babbie, 2010]. External validity refers to whether the causal relationships from the study can be generalized beyond the study conditions [Rubin and Babbie, 2010]. The validation activities conducted were done to achieve the following objectives: (1) to remove the potential threats to validity where possible and (2) to clearly understand the threats to validity where unavoidable. The value to clearly understanding the threats to validity where unavoidable is that these threats can be clearly documented along with the presentation of the survey results, and this reduces the risks of others misinterpreting or misapplying the findings from the survey.

C.2 Validation Activities for the Survey

The various validation activities that have occurred related to the survey are discussed here. This section discusses the validation activities related to the set of factors that are evaluated via the survey, validation related to the survey questions and validation of the survey instrument.

C.2.1 Validation of Set of Factors Assessed by Survey

The set of factors in the survey were selected because they are believed to be potential predictors of the likelihood of an organization being awarded a contract. However, this set was derived and refined via a number of steps involving a number of individuals. This section describes the process that was undertaken to establish and refine this set of factors.

The set of factors was initially captured in an analytical model that was part of the decision support system described in this dissertation. The set were the factors believed to be important enough that a decision maker may wish to focus on them in a proposal effort in an attempt to be awarded a contract. This initial set was selected based upon the experience and subject matter expertise of the author as well as a systems engineering strategy framework.

The analytical model was developed under the supervision of the dissertation supervisor, who also has experience with using systems engineering on proposals. The initial analytical model was reviewed with the other committee members in two meetings on February 22, 2012 and again on March 7, 2012. In these meetings, the factors were refined and a few additional factors were added. The analytical model was presented at the proposal defense on April 18, 2012. Throughout the process, the model was steadily refined.

The set of factors was also reviewed by members of the University of Texas at Arlington Systems Engineering Research Center (SERC) Industry Advisory Board (IAB) on June 29, 2012. The set of factors was presented in the form of a bullet list. Approximately 30 minutes was devoted to this topic. Various members shared opinions and observations related to the set of factors. As a result of the inputs, two factors were added: whether or not the customer organization's home office was in the United States or outside the United States and also the motivation of the organization for wishing to submit a proposal. Questions were added to gather this data.

C.2.2 Validation of Survey Questions

In addition to the need to determine if the correct set of factors is being measured by the survey, it is also necessary to validate that the questions are written effectively and clearly elicit the information that is intended. Therefore, separate validation activities were conducted for the questions. Note that some of the validation activities intended to validate the actual survey instrument actually worked as additional validation of the questions as well.

All committee members also reviewed the questions and provided feedback. This took place in April of 2012. The questions were also reviewed by two additional subject matter experts in April – May of 2012. Modifications were made as a result of each review. The questions were reviewed again in conjunction with peer testing and field testing.

C.2.3 Validation of Survey Instrument

The survey instrument was validated by field testers and expert reviewers. Field testers input real data into the survey for all required fields and provided feedback about any issues that they encountered as a result of populating the survey. Expert reviewers are people with the knowledge base representative of the core survey respondent demographic, but who are unable to input the required data. For the case study survey, five people were approached to field test. Three actually input data into the survey form and all five provided feedback. Of the three who input data, one entry was sufficiently complete to be considered a field test.

Appendix D

Survey Overview

D.1 Overview

This chapter is designed to provide an overview of information related to the survey. This includes the objectives of the survey, the survey target population, details of the administration of the survey and the demographics of the respondents, contracts and proposal efforts reported. This also involves a discussion of the general analysis strategy and the specific analysis techniques used. Detailed results of the survey are presented, both from an individual factor analysis and from an analysis that considers multiple factors simultaneously.

D.2 Survey Objectives

The survey was conducted to fulfill several research objectives. These include obtaining an understanding of how systems engineering is used in proposal management in practice, determining which systems engineering related factors appear to be have a statistically significant correlation with an increased probability of contract award and providing a set of statistical relationships that can be used to develop the DSS to find an optimal strategy for using systems engineering in the proposal process.

D.2.1 Obtain an Understanding of How Organizations are Using Systems Engineering in Proposal Management

Systems engineering is routinely used on proposal efforts to perform a number of tasks including align market opportunities with an organization's competencies and to define the solution that will be offered [Smartt and Ferreira, 2012c]. No research has been found that provides any quantitative description as to how systems engineering is being used on proposals and how much systems engineering is being done to define the proposed offering on a proposal. The results from *Survey to Assess the Relationships Between Systems Engineering Factors and Proposal Success* provides insight into how organizations are using systems engineering on proposals. While this information alone is not prescriptive in nature and does not provide direct guidance to organizations about how to apply systems engineering to improve their success with proposals, it does allow organizations to compare and contrast their practices with other organizations that are successful and unsuccessful in securing contract awards. For example, someone may determine from analyzing the data that their organization is an outlier in that it spends considerably more or less on systems engineering on proposals than other organizations that have been successful at obtaining contract awards. Through looking at results of the survey, someone may gain some understanding of the potential implications for contract awards.

D.2.2 Determine Which of the Systems Engineering Related Factors are Statistically Significant

Someone looking to use systems engineering to improve their level of success with proposals is likely to want to know exactly what factors to focus on. The results of this survey provide insight into which factors are correlated with an increased probability of contract award. People can then focus on these factors. A word of caution is in order: just because this study does not find a particular factor to be statistically significant does not mean that the factor should be ignored. For example, while the results do not show that holding a high capability maturity rating is necessarily correlated with an increased probability of contract award, holding a particular rating may be a prerequisite to enter a proposal for certain contract opportunities. If an organization wishes to be awarded such contracts, they must focus on process maturity.

D.2.3 Provide Statistical Data to Support the DSS

One major objective of the survey is to gain insight about relationships between variables that are used in example decision support systems that demonstrate the modeling methodology described in this dissertation. While the quantity of data collected by the survey is insufficient to derive parameters for a single DSS that can be used across organizations to yield valid recommendations for optimizing the use of systems engineering on proposals, the survey data does provide insight into key relationships between systems engineering factors on proposals and contract awards. The relationships derived from this data can be used to evaluate the realism of example models created for illustrative purposes.

D.3 Survey Target Population

The survey target population are people who have knowledge about how systems engineering was used on proposals. These include the stakeholders that are the target demographic for the DSS such as a functional systems engineering manager, a lead engineer, a lead systems engineer, a project manager and a program manager. Other types of people with sufficient knowledge to answer the questions may be marketing professions or capture managers, individual contributors, directors or executives who have insight into one or more proposal efforts and how systems engineering was used on the proposals.

D.4 Survey Administration Details

This section provides details about the survey administration that may be relevant for interpreting the survey results.

D.4.1 Development Environment and Implementation Characteristics

Two versions of the survey were ultimately created and deployed. The first version of the survey was coded using Microsoft Expressions and hosted on a domain supported by a commercial host site. The Microsoft Expressions version of the survey was largely unsuccessful at eliciting responses. After a careful analysis, it was determined that many of the potential respondents work for organizations who implement very stringent controls on employees' use of the internet, and in many cases block employee access to sites that have not explicitly been pre-approved, such as the commercial service hosting the survey. As a result, the survey had to be re-hosted using a commercial off the shelf (OTS) tool called Survey Monkey. An analysis revealed Survey Monkey to be accessible from within the firewalls of several organizations that employ potential respondents. Therefore, this section provides implementation details about each version of the survey.

The Microsoft Expressions version of the survey was coded using three custom web pages. The first page was a welcome page that included the following information: the survey purpose, information about the confidentiality of the respondents, the uses that would be made of the data, the qualifications of potential respondents, incentives, how results would be reported, and what hardware and software is recommended for the best results. The second page was the survey. On this page, the respondents filled out the entire survey, and then selected the submit button. Front Page server extensions were then used to capture the data that has been entered into the survey form and record it to a database. The software also sent an email reporting the data entry as a backup. The third page was simply a thank you page that informed the respondent that he or she had been entered into a drawing for one of three Amazon.com gift certificates. There was no way in the Microsoft Expressions version of the survey to save a partial entry and return later. Therefore, a PDF version of the survey was offered for respondents so that they could gather the necessary information prior to filling the survey out.

The Survey Monkey version of the survey was coded as 11 separate pages. This larger number of pages was necessary because Survey Monkey offers the option to code question logic into the survey, and the questions partition the survey because certain questions are not presented to individuals who provided particular answers to other questions. Question logic was included in the *Survey to Assess the Relationships Between Systems Engineering Factors and Proposal Success*. The question logic feature proved useful because it allowed unqualified respondents to be filtered and ejected from the survey early, and it allowed qualified respondents to be automatically routed to only the relevant questions. The Survey Monkey version of the survey is provided in Appendix A.

The directions for both the Microsoft Expressions version of the survey and the Survey Monkey version of the survey specifically indicated that input was only requested for proposal efforts for which a definitive award status is known. If it was unknown whether a contract was awarded or the final outcome was pending, then potential respondents were instructed not to provide input related to that proposal effort. Nonetheless, sixteen people proceeded to begin the survey who were reporting on proposal efforts for which the final award status was unknown. The question as to the contract award status was asked early in the questionnaire. Respondents who indicated that the award status was unknown were ejected from the survey. This was a useful feature because it avoided additional time being wasted on the part of the respondents.

For qualified respondents reporting on proposal efforts within the scope of the research, not all questions were necessarily relevant. For example, one question asked about the process maturity of the prime contractor. A previous question asked the respondent to indicate whether their employer was the prime contractor or a subcontractor, and another previous question had asked the respondent what the process maturity level of their organization was. If the respondent answered that they worked for the prime contractor and had reported the process maturity level of their employer, by definition they had provided the process maturity of the prime contractor and no additional question was necessary.

D.4.2 Timeframe

Field testing occurred between August 5, 2012 and August 30, 2012. The Microsoft Expressions survey collected data from September 21, 2012 to October 15, 2012. The Survey Monkey version of the survey collected data from November 13, 2012 until December 15, 2012.

D.4.3 Responses and Drop Outs

Table D.1 provides a summary of the responses received to the survey from field testers as well as responses to both the Microsoft Expressions and the Survey Monkey versions of the survey. The remainder of this section will provide a more detailed explanation for these sources.

Table D.1 Summary of Responses

Number field testers	3
Number of individuals who completed Microsoft Expressions survey	4
Number of individuals who completed the Survey Monkey version of the survey	61
Total number of individuals who completed the survey	68
Number of responses excluded	6
Number of valid responses analyzed	62

D.4.4 Access Statistics for Microsoft Expressions Version of the Survey

An exact number of hits for the cover of the Microsoft Expressions version of the survey is unknown because the information for October is for the entire month of October, not just for the period prior to October 15 which was the valid window to enter data. On the other hand, the data for September was reset prior to the survey going live. Therefore, the statistics in Table D.2 are presented for the timeframe from September 21, 2012 – October 31, 2012.

Table D.2 Statistics for Microsoft Expressions Version of Survey

Number of times cover page was accessed	212
Number of times actual survey web page was accessed	52
Number of complete entries	4
Number of people who completed the survey whose responses were assessed to be valid entries	3

It appears that there is a dropout rate of over 98%. Less than 25% of people who read the cover proceeded to the actual survey, and just over 6% of the people who proceeded to the survey submitted results.

D.4.5 Access Statistics for Survey Monkey Version of Survey

The types of access statistics that are available through Survey Monkey are different than the types of statistics available from the commercial web service that hosted the Microsoft Expressions version of the survey. Specifically, Survey Monkey only reports respondents who "started the survey". Survey Monkey defines "started the survey" as having either submitted data or proceeded to the second page of the survey. For the case of the *Survey to Assess the Relationships Between Systems Engineering Factors and Proposal Success*, there was no information to be submitted on the first page as the first page was exclusively information about the purpose of the survey, the confidentiality of respondents, the distribution of incentives and recommended hardware and software for taking the survey. Survey Monkey does not indicate how many people may have viewed the first page of the survey and then did not proceed. Available statistics for the Survey Monkey survey for the timeframe from November 13, 2012 to December 15, 2012 are as listed in Table D.3.

D.4.6 Drop Out Rates

The response rates and drop outs varied radically between the Microsoft Expressions version of the survey and the Survey Monkey version of the survey. A potential cause for this discrepancy is how the survey was represented. INCOSE placed a direct link to both versions of the survey on their home page. The link that was posted to the Microsoft Expressions version read “Do you have 30 minutes to complete a PhD survey?” The link to the Survey Monkey version read “Your Help is Requested on a Systems Engineering Survey Related to Factors for Winning Proposals”.   Because the second prompt provided information as to the content of the survey, people who clicked the link likely had interest and/or knowledge of the subject matter of the survey.

Table D.3 Statistics for Survey Monkey Version of Survey

Number of people who began survey	127
Number of people who were ejected from the survey because they indicated that the contract award status as a result of the proposal was unknown	16
Number of people who completed the survey	61
Number of people who completed the survey whose responses were assessed to be valid entries	56

D.4.7 Data Validation

In Table D.2 and Table D.3 there are separate entries for the number of people who completed the survey and the number of people whose responses were assessed to be valid entries. It is important to validate responses and exclude invalid responses from the analysis so that the invalid responses do not distort the results and potentially even the findings of the survey analysis. Mistaken inclusion of invalid data could lead to incorrect evaluation of the hypotheses in the analysis. For validation, a number of questions were asked to help ensure that the entries represented valid data. This was done to protect against invalid data affecting the various statistical relationships that are derived from this data. The following fields of data were collected for validation purposes: (1) respondent’s name, (2) respondent’s email address or telephone number, (3) a unique identifier for the project, (4) the total number of years of experience for the respondents, (5) the number of years of experience for the respondent separated into a number of different roles, including a role for “other”, (6) a question asking respondents to report their level of confidence in the qualitative inputs provided, and (7) a question asking respondents to report their level of confidence in the quantitative inputs provided.

Responses can be excluded from particular parts of the analysis. For example, some responses are excluded from the categorical factor analysis, the effort data quantitative analysis and the analysis of the effects of the number of customer interactions. The categorical analysis includes the demographic analysis of the respondents, the organizations that employee the respondents, the contracts and the proposal effort, and the factor-by-factor analysis of the effects of various variables that are categorical in nature on the probability of contract award. The categorical analysis does not include questions where the factor being analyzed is on a continuous scale (e.g., a ratio). The effort data quantitative analysis pertains to the effect of the relative (to contract size) level of investment in systems engineering labor hours to define the proposed offering on the probability of contract award. The analysis of the effects of number of customer interactions on the probability of contract award analyzes the effects of the number of various different types of interactions (e.g., written, telephone, video conference, in-person, total) relative to the contract size on the probability of the contract award. All responses excluded from the categorical analysis are automatically excluded from the effort data quantitative analysis and automatically excluded from the number of customer interactions analysis.

A response is excluded from the entire analysis (not considered a valid entry) if: (1) the name is missing, (2) neither the email nor phone number is provided, or (3) the respondent indicates that he or she is "not confident at all" related to either qualitative or quantitative responses. A number of respondents elected not to provide a unique identifier for the project. It is plausible that someone who is providing valid entries may elect not to supply this information because they felt disclosing this information could place an organization at risk and did not elect to provide an alternative name. Therefore, respondents who elected not to supply this information did not have their inputs automatically excluded. However, if the unique project identifier was not provided and there was a discrepancy between their reported total years of experience and the sum of their years of experience in various roles, then the entry was excluded. Some responses that passed the data checks and were included in the general analysis may still be excluded from particular parts of the analysis.

There were a total of 62 valid survey responses included in the categorical data analysis. Of those 62 responses, only 37 of the responses were considered usable for the effort data quantitative analysis. The effort data quantitative analysis relates to exploring the significance of investing labor hours in core systems engineering activities while defining the proposed offering in the proposal and the significance of the number of customer interactions. There are a number of reasons certain responses were excluded from this particular portion of the analysis. For example, a number of respondents did not provide estimates of the number of systems engineering labor hours to be awarded if a contract was awarded or the total number of engineering hours (including systems engineering) to be awarded if the contract was awarded. Because these two variables are used to size the raw number of labor hours spent defining the proposed offering so that proposal efforts for large contracts can be compared with proposal efforts for smaller contracts. Another reason certain responses were excluded is that the respondent reported no effort for all of the five core activities and for all of the four employee skill levels. In these cases, the respondents did not enter any values in any of the fields corresponding to the number of labor hours for activities or skill levels. Still other survey responses were excluded from the analysis because the respondent reported that there were more systems engineering hours to be awarded on the contract than there were total number of engineering hours (including systems engineering hours).

D.5 Demographics

This section explores the demographics of the respondents, their organizations, the prime contractor organizations, the types of systems for which the contracts where supporting, the characteristics of the contracts and the characteristics of the proposal effort. The purpose of analyzing and reporting the demographics is so that readers understand the context of the data collected and the context of the analysis presented. Respondents will know how applicable the survey findings and optimization model are to their particular problems and can make informed decisions about whether to follow or not follow any recommendations derived from the survey analysis.

D.5.1 Proposal Effort Project Characteristics

Figure D.1 displays the year various proposals were submitted. Only valid data is displayed.

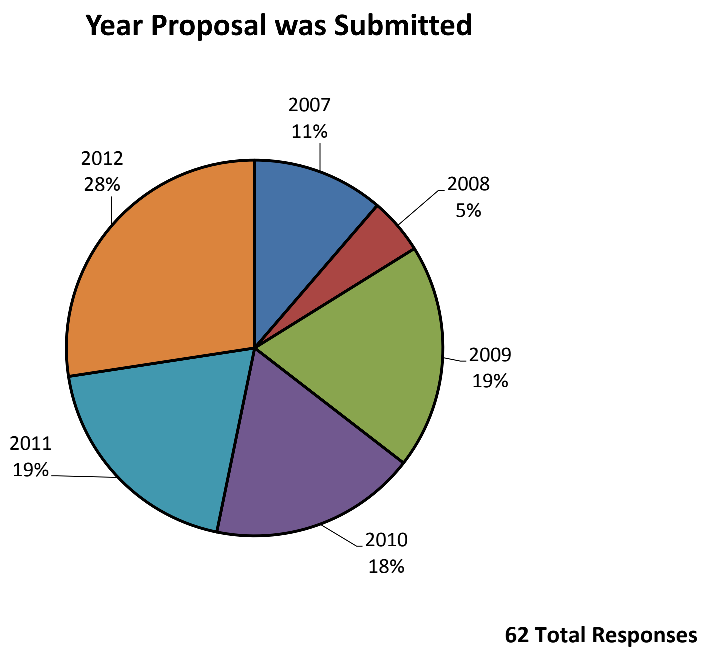


Figure D.1 Year Proposal was Submitted

These years correspond only to the data points used in the statistical analysis. One respondent reported on a proposal that was submitted in the 1990s. Because the directions in the survey explicitly directed respondents to reply only if the proposal was submitted within the last five years (which would have be 2007-2012 per the dates of the survey), the proposal submitted in the 1990s was not included in the analysis.

D.5.2 Survey Respondent Characteristics

This section describes characteristics of the survey respondents that may be useful for interpreting the results of the survey analysis. Figure D.2 displays the professional organization affiliations of the survey respondents.

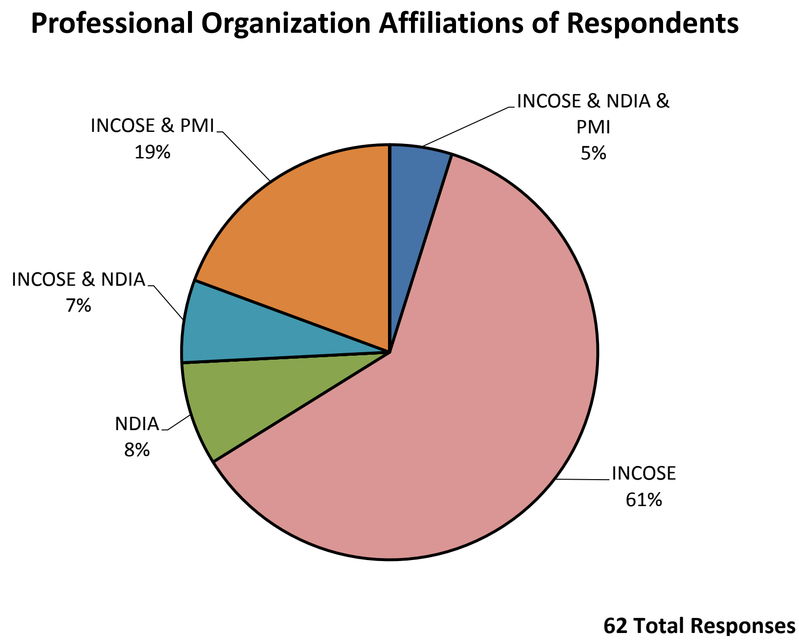


Figure D.2 Professional Organization Affiliations and Respondents

Respondents were given an opportunity to indicate whether they belonged to the International Council on Systems Engineering (INCOSE), the National Defense Industrial Association (NDIA) and the Project Management Institute (PMI). Respondents could select none up to all of these organizations. Of the 62 respondents, 92% were INCOSE members at the time that they responded to the survey. Many of the INCOSE members were also members of other professional organizations. The fact that all of the survey respondents belonged to professional organizations has potential implications for how the survey results can be generalized across industry. This raises a question of whether the survey respondents are representative of professionals in general who use systems engineering to define the proposed offering on a proposal. It is unknown to what degree this is an issue.

Figure D.3 displays the respondents' experience by role.

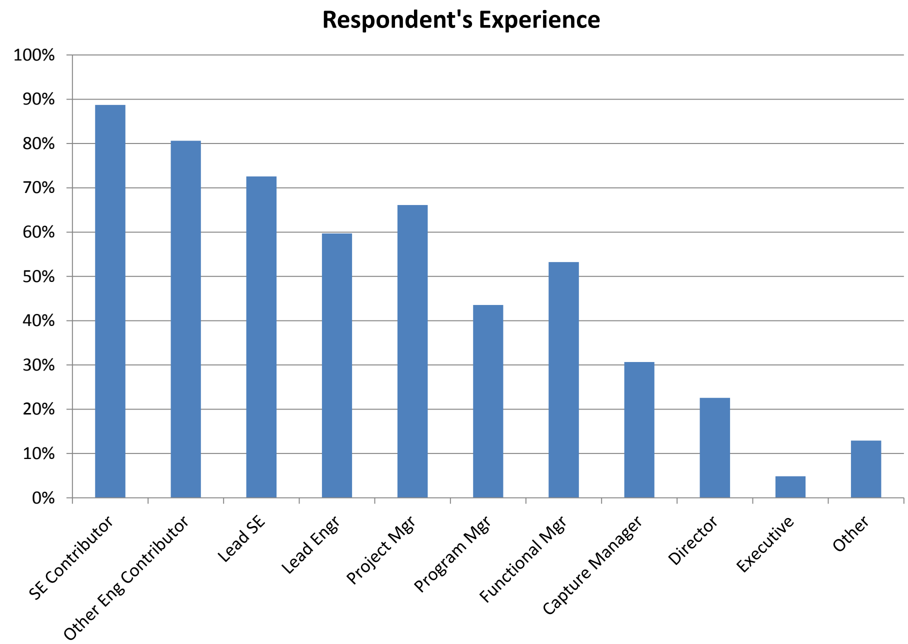


Figure D.3 Respondents' Experience by Role

Almost 90% of the respondents have at least some experience as systems engineering contributors. Over 70% have experience as a lead systems engineer. Over 60% have experience as a project manager, and almost 60% have experience as a lead engineer or a functional manager of a group of individuals performing systems engineering. Several respondents had experience as a capture manager, director or executive. The implication of the respondent experiences is that many respondents had experience in a number of different types of roles.

Figure D.4 displays the respondents' role or roles on the proposal effort that they were reporting on.

More than half of the respondents were systems engineering contributors on the proposal effort, over 25% of respondents were lead systems engineers, and more than 20% of respondents were project managers. There was at least one respondent who served in each of the designated roles.

Figure D.5 displays the average number of years of experience in each role.

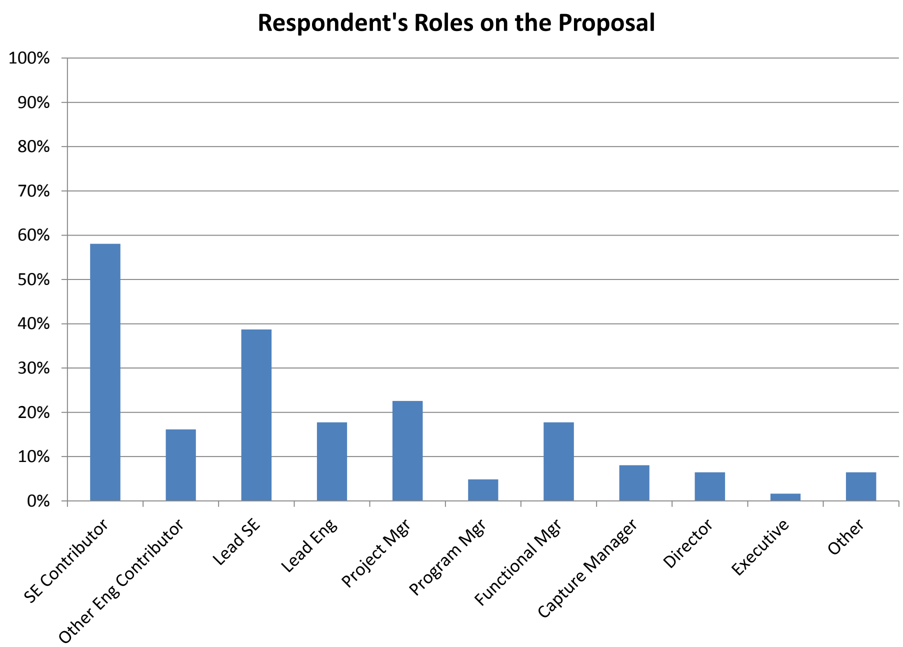


Figure D.4 Percentage of Responses for Each Role on the Proposal

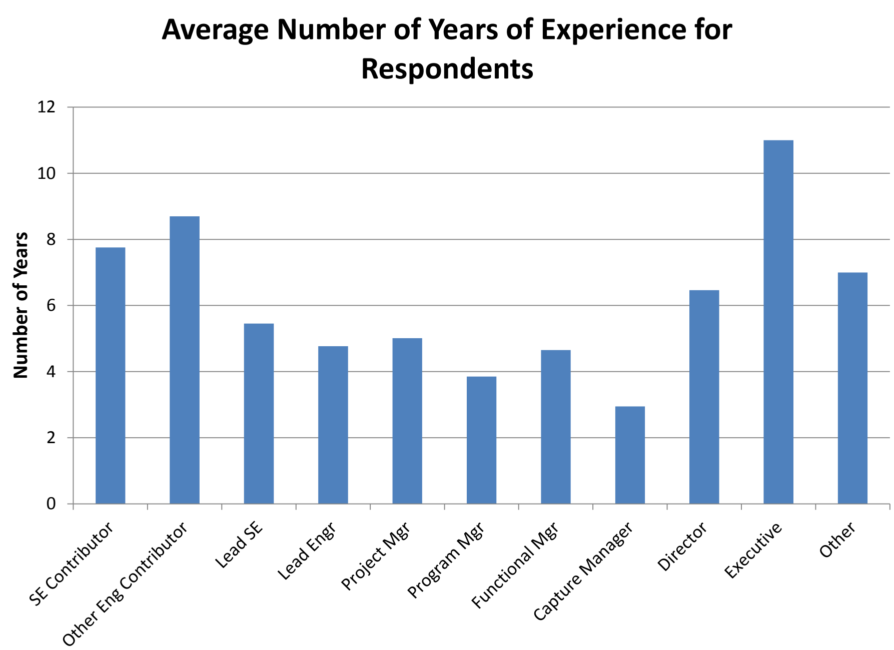


Figure D.5 Average Number of Years of Experience for Respondents

The average number of years of experience for respondents is derived for the people who indicated that they had experience in a particular category in Figure D.3. For example, of the 88% of respondents who indicated that they had some experience as a systems engineering contributor in Figure D.3, the average number of years of that experience is about 7.7 years. This indicates that not only were respondents experienced in a number of different roles, but that generally speaking the respondents had a considerable number of years of experience in those roles.

Figure D.6 describes the level of confidence that respondents have in their qualitative inputs or answers to categorical questions.

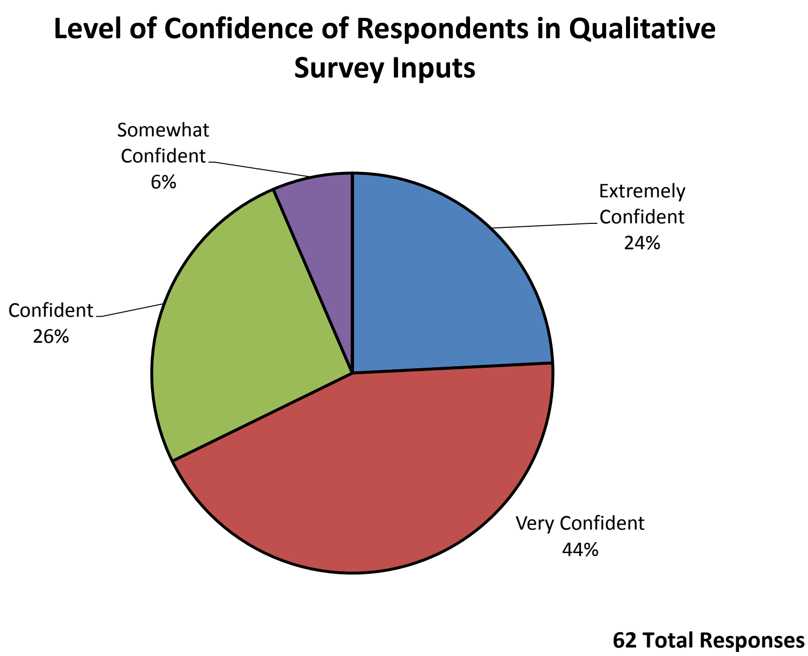


Figure D.6 Level of Confidence of Respondents in Qualitative Survey Inputs

94% of respondents indicated that they were confident, very confident or extremely confident in their responses. 6% indicated that they were somewhat confident. No respondents indicated that they were not at all confident.

Figure D.7 describes the level of confidence that respondents have in the answers that they provided for quantitative questions.

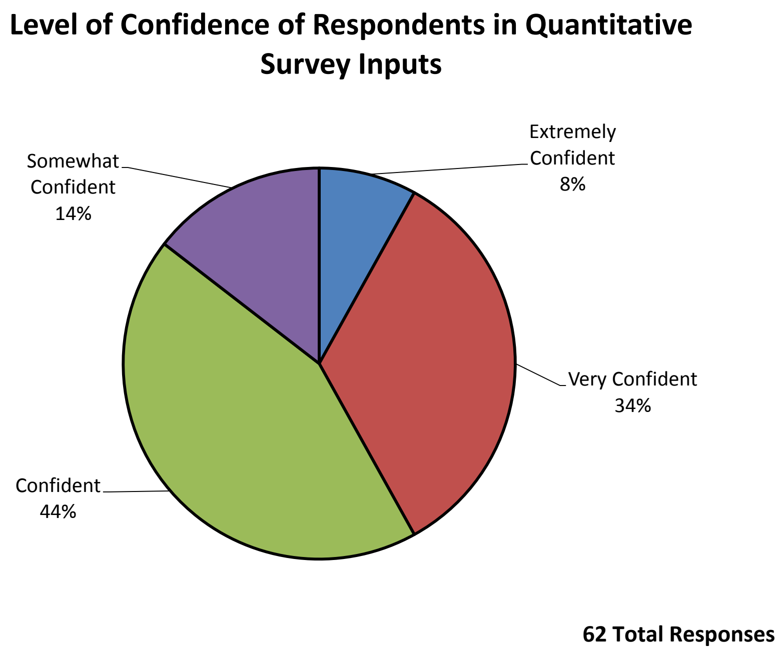


Figure D.7 Level of Confidence of Respondents in Quantitative Survey Inputs

86% of respondents indicate that they were confident, very confident or extremely confident in their answers to quantitative questions. 14% of respondents indicated that they were somewhat confident in their responses. No respondents indicated that they were not confident at all in their responses. While these confidence levels are slightly lower than the respondents' confidence levels for qualitative questions reported in Figure D.6, as a whole this indicates that respondents still have a fairly high level of confidence in their quantitative inputs.

The high level of confidence of respondents in both their qualitative and quantitative inputs is encouraging. However, this is no guarantee that the data collected by the survey accurately reflects reality for the proposals described. It is possible that respondents in some cases may have unknowingly misunderstood the questions or failed to read important explanatory information meant to guide their responses and then confidently provided incorrect answers.

D.5.3 Supplier Organization Characteristics

Of the organizations that employed respondents of the survey, 79.4% served in the role of prime contractor and 20.6% served in the role of subcontractor. In the survey, a prime contractor is defined as an organization that contracts directly with the government. A subcontractor is defined as an organization that is not a prime contractor. All 62 survey respondents answered this question.

D.5.4 Acquirer Organization Characteristics

Of the acquirer organizations that were the customers who the proposals were written for, 63.5% had home offices located in the United States and 36.5% had home offices located outside the United States. All 62 survey respondents answered this question.

D.5.5 System Characteristics

This section describes characteristics of the types of systems reported by survey respondents. Survey respondents were allowed to select zero up to all of the following categories: banking/financial systems, communications systems, electronics systems, healthcare systems, information systems, infrastructure systems, shipping/distribution/logistics systems, training systems, transportation systems, weapon systems and other types of systems. Figure D.8 reports the number of different types of systems respondents selected.

All respondents selected at least one type of system. 79% of respondents selected exactly one type of system. 14% of respondents selected two types of systems, and 7% of respondents selected three or more different types of systems.

Figure D.9 provides a display of number of survey responses for which respondents selected a particular system category.

Almost half of the proposals were for system types that included weapons systems. A significant number of proposals were for information systems and transportation systems. Figure D.10 displays the number of proposals by sector.

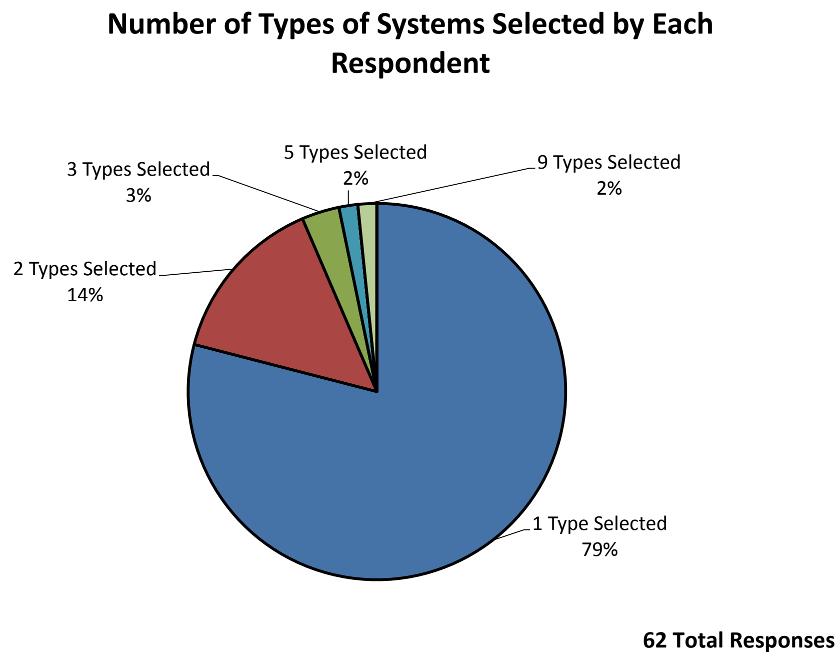


Figure D.8 Number of Types of Systems Selected by Each Respondent

Respondents were allowed to select none up to all three of the following categories: commercial, military/defense and government/non-military. 44 of the 62 respondents indicated military/defense, 10 respondents selected commercial and 10 respondents selected government/non-military.

All 62 respondents indicated whether the system to be developed under the contract was first of a kind, a replacement of or an enhancement to an existing system. 31% indicated first of its kind, 27% indicated an enhancement to an existing system and 42% indicated a replacement to an existing system.

D.5.6 Contract Characteristics

Figure D.11 displays the types of contracts that are being pursued in the proposal efforts. 63% of respondents reported firm fixed price contracts, 8% cost plus percent cost, 18% cost plus fixed fee and 11% some other type of arrangement. All 62 respondents answered this question and provided exactly one answer.

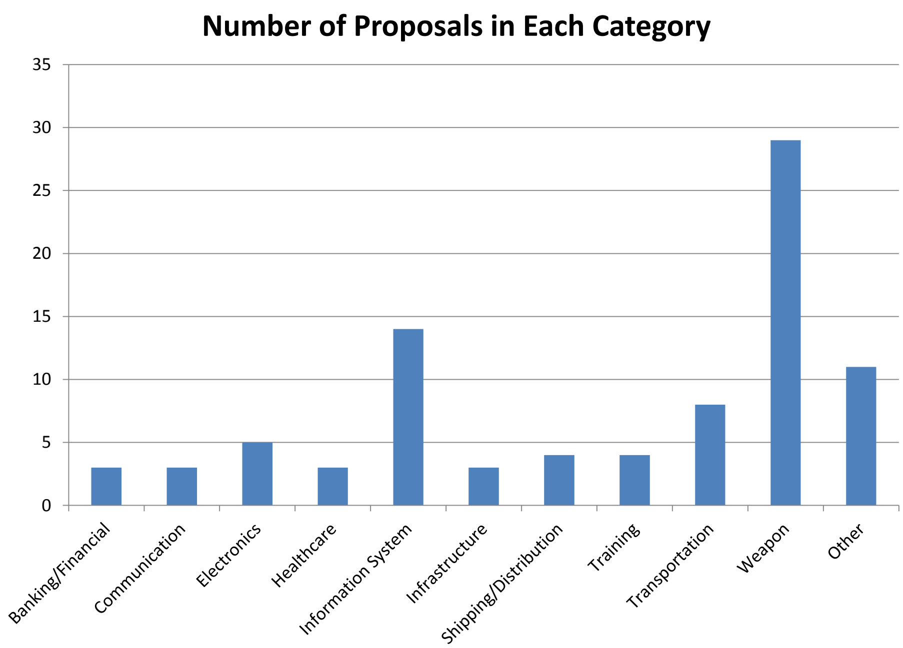


Figure D.9 Number of Proposals By Category

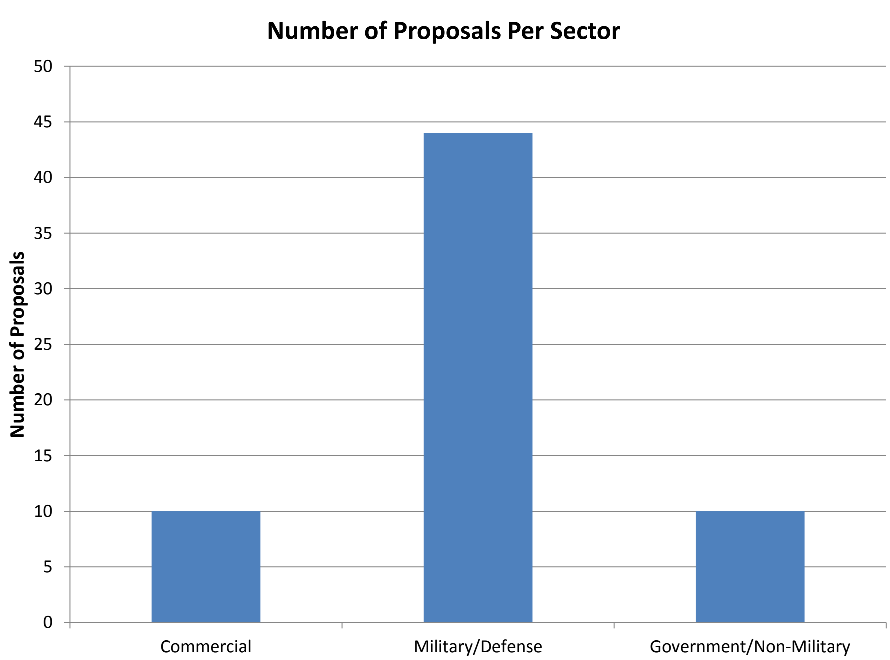


Figure D.10 Number of Proposals Per Sector

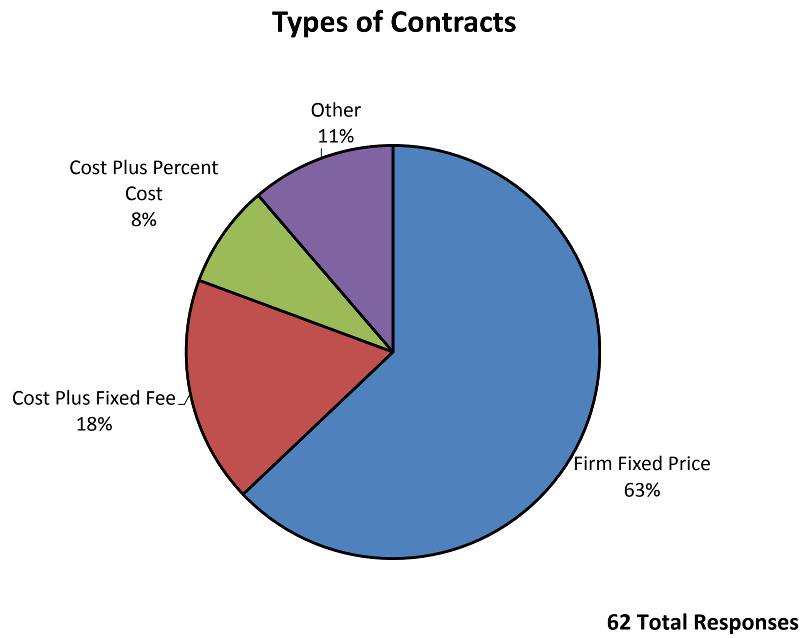


Figure D.11 Types of Contracts

D.6 Definition of Factors Analyzed

This section defines the factors that are analyzed. In later sections of this appendix and in the body of the dissertation, these factors are referred to and statistics are presented related to these factors. Two types of factors are defined, continuous and categorical. The continuous factors relate to the number of contacts between the prime contractor and the customer and also the effort on the contract. All continuous factors are normalized. Normalization is defined as the actual raw counts of the number of contacts and number of labor hours spent defining the proposed offering divided by the total number of engineering labor hours on the development portion of the contract. Normalizing is necessary because the range in the sizes of the contracts spans several orders of magnitude. Specifically, the size of the development portion of the contracts varies from less than 100 labor hours to millions of labor hours. The number of systems engineering labor hours spent defining the proposed offering on the proposal divided by the total number of engineering labor hours to be spent on the development portion of the contract is referred to throughout the analysis as DPORatio. Some portions of this, related to certain activities or particular skill levels, are shown with a dash after DPORatio. For example, the number of systems engineering labor hours spent on source requirements divided by the total number of engineering labor hours to be spent on the development portion of the contract, is referred to as DPORatio – Source Requirements.

Categorical factors have a finite number of levels such as “no”/”yes” or levels 1-5. Table D.4 defines the categorical factors analyzed. For each factor, a name, definition and list of levels are provided. To supply additional context, the numbers for the question or questions in the survey provided in Appendix A used to collect data related to each factor is provided.

D.7 General Analysis Strategy

This section describes the general analysis strategy including the approach that will be applied in analysis and model building as well as the strategy for forming hypotheses. A number of techniques for model building that are applied in this research can be found in standard regression textbooks. These techniques help select the subset of covariates that should be included in a model as well as any interaction terms that are significant enough to warrant including in the model. Examples of such techniques include the best subsets algorithm, backwards elimination, forward selection and forward stepwise regression [Kutner et al., 2005; Montgomery et al., 2012]. From any subset of p-1 predictors, 2p-1 alternative models can be constructed [Kutner et al., 2005]. The survey examines 16 categorical predictors and 25 total continuous predictors. As opposed to relying on an automated process (e.g., best subsets) to evaluate all of these model combinations, some preliminary factor-by-factor analysis is done first to reduce the number of parameters to even consider. Because no quantitative work has been done previously to examine these factors, it is expected that just a few of the factors will hold any explanatory power at all, and most will not.

As a first step, factors are examined individually to see if they have at least some minor degree (e.g., p value > 0.2) of significance. ρ-values are the smallest value of alpha (i.e., significance level), that would allow the null hypothesis to be rejected [Dean and Voss, 1999]. ρ-values are between 0 and 1 by definition. ρ-values less than 0.05 are considered significant because 0.05 is a commonly used threshold in statistical analysis. ρ-values between 0.05 and 0.2 are determined in this analysis to be mildly significant. This means that while these factors cannot by traditional standards be considered significant, it is quite possible that with a larger sample size these factors would indeed be determined to be statistically significant. Because of the relatively small number of survey responses, all factors that are not determined to be significant or mildly significant are inconclusive. With the sample sizes in the data, there is no way that the null hypotheses (see next section) can be proven. In these cases where the ρ-values exceed 0.2 for all of the statistical tests, the way to interpret the results is that the null hypothesis cannot be rejected. It does not mean that the null hypothesis is accepted. For each factor in each statistical model, great care is placed in estimating the level of significance. For estimates that rely upon conformity of the data to some underlying model (e.g., Wald), those model assumptions are examined. The data does not conform to the model assumptions and therefore alternative methods for estimating the significance of the factors are explored.

As a second step, more traditional model building approaches are examined using only this smaller subset of factors to build a single, unified multivariate regression model to be used as part of the optimization model. The smaller subset will include all factors with a ρ-value of 0.2 or less. The subset will also include any parameters for which inspection suggests that there is a relationship but the statistics are inconclusive. In this second process, the correlation or relationships between factors as well as the satisfaction of model assumptions is thoroughly explored.

The benefit to this is that the model search process for the multivariate regression model is more tractable than if every single data items collected had been considered. It is expected that with a small data set, such as the set of survey responses, that there may be computational issues building some of the model combinations (e.g., splitting), and therefore a significant amount of remediation and attention is applied in the model building process.

Table D.4 Definition of Categorical Factors and Factor Levels

Factor Name	Definition	Levels	Relevant Question Number(s)
Type of Contact	The type of contractual arrangement that was applicable or would have been applicable if a contract was awarded	a. Firm, fixed price b. Cost plus fixed fee c. Cost plus percentage cost d. Other type of arrangement	13
Motive - Profit from Development	The organization submitted a proposal in order to make a profit from executing the development portion of the contract	a. Yes b. No	15
Motive - Profit After Development	The organization submitted a proposal in order to make a profit from operating, maintaining or disposing of the system after delivery	a. Yes b. No	15
Motive - Future Work	The organization submitted a proposal in order to increase the probability of being considered for future opportunities involving the same customer	a. Yes b. No	15
Motive - Future Relationships	The organization submitted a proposal in order to establish a business relationship with other supplier organizations	a. Yes b. No	15
Motive - Satisfy Existing Customer	The organization submitted a proposal in order to satisfy an existing customer who expected a response	a. Yes b. No	15
Motive - Block Competitor	The organization submitted a proposal in order to deny a potential competitor an opportunity	a. Yes b. No	15
Systems and Software Process Maturity of Respondent Organization	Process maturity rating of the organization employing respondent, using levels from the staged representation of CMMI for Development v1.3 dated November, 2010	a. CMMI Level 1: Initial b. CMMI Level 2: Managed c. CMMI Level 3: Defined d. CMMI Level 4: Quantitatively Managed e. CMMI Level 5: Optimizing	25
System and	Process maturity rating of the proposal	a. CMMI Level 1: Initial	26

Table D.4—*Continued*

Software Process Maturity on Project	project, using levels from the staged representation of CMMI for Development v1.3 dated November, 2010	b. CMMI Level 2: Managed c. CMMI Level 3: Defined d. CMMI Level 4: Quantitatively Managed e. CMMI Level 5: Optimizing	
System and Software Process Maturity of Prime	Process maturity rating of the prime organization, using levels from the staged representation of CMMI for Development v1.3 dated November, 2010	a. CMMI Level 1: Initial b. CMMI Level 2: Managed c. CMMI Level 3: Defined d. CMMI Level 4: Quantitatively Managed e. CMMI Level 5: Optimizing	28
Level of Access	The level of freedom the prime contractor had to interact with the customer during the proposal stage	a. No access to the customer b. Authorization to interact with the customer a set number of times as specified by the customer c. Authorization to interact with the customer whenever the prime felt the need to interact d. Other type of customer interaction	29
Level of Project Manager Experience	Extent of the experience of the project manager of the proposal project for the prime contractor	a. Led more than one proposal b. Led exactly one proposal effort c. Never been an overall lead (project manager) on a proposal effort prior to the proposal effort described here, but was highly experienced in proposal efforts as either a contributor or a leader other than the overall leader (e.g., chief systems engineer, IPT lead) d. Never been an overall lead (project manager) on a proposal effort prior to the proposal effort described here, but had some experience in proposal efforts as either a contributor or a leader other than the overall leader (e.g., chief systems engineer, IPT lead) e. Had no experience with proposal efforts prior to the proposal effort described here	32

Table D.4—*Continued*

		f. Unknown	
Rank - Market Share	Competitive rank of the prime contractor in revenue from sales	a. 75 – 100 percentile b. 50 – 74 percentile c. 25 – 49 percentile d. 0 - 24 percentile e. Unknown	33
Rank - Prestige	The competitive rank of the prime contractor in prestige	a. 75 – 100 percentile b. 50 – 74 percentile c. 25 – 49 percentile d. 0 - 24 percentile e. Unknown	34
Extent of Relationship with Proposal Customer	Extent of the previous relationship (including both contracts and contacts) between the prime contractor and the customer	a. Was executing a separate contract with the intended customer at the time the proposal effort was submitted b. Had executed a separate contract with the intended customer within a two year period prior to the submission of the proposal, but was not executing a contract with the customer at the time of the proposal submission c. Had no separate contracts with the customer within a two year period, but did have past contacts with the customer d. Has had no existing previous contacts or contracts with the intended customer	35
Level of Customer Satisfaction on Previous or Ongoing Effort	In past or ongoing contracts with the prime contractor at the time of the proposal submission, this customer's general level of satisfaction with the performance of the prime contractor	a. Very satisfied b. Somewhat satisfied c. Neither satisfied nor dissatisfied d. Somewhat dissatisfied e. Very dissatisfied	36

The disadvantage to doing a preliminary down-select of candidate factors is that a greater potential exists to incorrectly eliminate latent explanatory variables before the multi-variable regression model building process even begins. Latent explanatory variables may be part of an important interaction, but do not necessarily appear significant outside that interaction. Alternatively, latent explanatory variables may be highly significant within a limited range of possible values for some model parameter, but insignificant outside that range. This potential to eliminate latent explanatory variables is why a factor-by-factor inspection of the results is done in addition to statistical analysis.

For the categorical factors, there is not necessarily a directional hypotheses related to the outcome. For example, if a respondent indicates that his or her employer's organization pursued a contract as a means to make a profit supporting the system after development, the researcher has no a priori expectation of whether this should equate to a lower probability of contract award, a higher probability of contract award or no difference in probability of contract award. Therefore, for categorical factors, the hypotheses are formulated as follows:

H0: The probability of contract award does not vary as the level of the independent variable varies.

H1: The probability of contract award varies as the level of the independent variable varies.

For cases where it appears plausible that H0 should be rejected, further post-hoc analysis is undertaken to understand the nature of how the probability of contract award varies as a function of the different levels of the categorical variables. Post-hoc multiple comparison methods are used (e.g., Tukey and Scheffé). The formulas for these are adjusted for "messy data", such as comparing categories where the number of data points are different [Milliken and Johnson, 1997].

For the continuous factors, there is some expectation as to the direction of the hypotheses. For example, investing more in systems engineering labor defining the proposed offering should lead to a higher probability of contract award. Likewise, more interactions with the customer (e.g., in-person interactions) should lead to a higher probability of contract award. Therefore, for continuous factors, the hypotheses take the following form:

H0: Increasing the value of the factor does not result in a higher probability of contract award.

H1: Increasing the value of the factor does result in a higher probability of contract award.

The remaining single factor statistical analysis is designed to evaluate these hypotheses and their level of significance.

D.8 Techniques Used for Individual Factor Analysis

This section discusses the various techniques used in the individual factor statistical analysis. This is a factor-by-factor exploration of statistical significance.

D.8.1 Techniques for Evaluating Categorical Factors

The most straightforward approach to evaluating the various hypotheses related to the significance of parameters is by assuming the data conforms to some well-understood, asymptotic model and using established tests related to that model to estimate the significance of the parameter. The results of such tests depend upon the data. In some cases, a certain quantity of data is necessary. In other cases, satisfaction of certain model assumptions is necessary.

For categorical factors, two asymptotic tests were performed. One test was a one-way analysis of variance [Mathworks, 2013]. One-way analysis of variance compares means of two or more sample groups. In the analysis of the survey results, the sample groups corresponded to the levels of the factor being analyzed. One-way ANOVA does have some implicit assumptions, including constant variance between the groups and normally distributed data. The other asymptotic test that was performed was a Pearson Chi-Squared Test in SPSS [IBM, 2013]. Pearson's Chi-Squared Test examines a null hypothesis that the frequency of a certain outcome is equally likely in each of the groups being examined. For the Pearson test, the groups were the various levels of each of the categorical factors.

For some tests, the expected frequency is very small. For example, for maturity level of project, only two respondents selected “Level 5”. Under the null hypothesis that award status does not vary as the maturity level on project varies, one would expect there to be 19/62*2 = 0.61 respondents who reported Level 5 to not be awarded a contract. By rules of thumb developed by Cochran [1954], for tables larger than 2 x 2, a minimum expected count of 1 is permissible as long as no more than about 20% of the cells have expected values below 5. Clearly, Cochran’s rule is violated for many of the categorical tests, including the example mentioned as 0.61 is less than one.

An alternative to these tests are exact tests. Exact tests make no assumptions and generate estimates only from the data provided [Mehta and Patel, 2011]. The exact Pearson Chi-Square test estimates significance based upon the likelihood of the given observation among the exhaustive set of possible observations. As a result, for even medium data sets, exact calculations are computationally intensive. However, exact tests are more computationally feasible for situations with binary outcomes. SPSS is the software used for exact computations for the categorical factors.

Post-hoc comparisons (Tukey and Scheffé) were performed based upon the results of the individual factor tests for factors for which ANOVA, Pearson’s Chi-square test or Exact Pearson’s Chi-square indicate a ρ-value of 0.2 or less. This is done to pinpoint which set of factors are statistically different than which other set of factors. Contrasts were defined to compare various sets of factor levels with other sets of factor levels.

D.8.2 Techniques for Evaluating Continuous Factors

For cases where the factor is a continuous variable (e.g., ratio of number of systems engineering hours spent defining the proposed offering on the proposal to the number of total engineering hours to be awarded if the contract is awarded) and the outcome is binary (award, no award), the logistic regression model is a well-accepted and well-understood model to use to make predictions and evaluate hypotheses (see Section 2.2.3.2). While many of the same rigorous tests do not exist for logistic regression model assumption satisfaction that exist for linear regression models (e.g., Modified Levene), generally normality, constant variance and the absence of extreme outliers are considered indicators that a logistic regression model is likely appropriate.

The main objective of the single factor analysis is to evaluate whether or not each factor is statistically correlated with contract awards. This equates to a hypothesis involving a model parameter or alternatively forming a confidence interval (CI) about a parameter of interest and seeing if the null value (i.e., 0) is included in the CI. The corresponding decision process is to reject H0 only if the null value is not in the CI.

The asymptotic tests most commonly used to make inferences for large sets of data for which a logistic regression model is appropriate are Wald tests [Kutner et al., 2005]. For asymptotic tests to be used appropriately, the data set must be sufficiently large. To verify that an adequate sample size exists, bootstrapping is used. Bootstrapping is recommended as a way to determine how dependent the model is on particular data points in the set [Kutner et al., 2005]. Bootstrapping is a Monte Carlo approach where in each iteration *M* data points are drawn randomly from the original data points (assuming there are *M* original data points) with replacement. In other words, for each iteration, some data points are used multiple times as if the same data point occurred multiple times while other data points are not used. For each iteration estimates for model parameters are calculated. After *N* such instances of model parameters are calculated from *N* iterations, the variation in the point estimates of the parameters are quantified. If the set of point estimates are symmetrically distributed about the mean of the point estimate from the actual data set and the standard deviation of the *N* point estimates is approximately equal to the estimated standard error for the parameter from the model with the original data set, then it is generally concluded that the data set is sufficiently large to use the asymptotic results from the logistic regression model. If this is not the case, confidence intervals about the parameters of interest can be estimated based upon the distribution of bootstrapped point estimates.

There are a number of different ways statisticians estimate confidence intervals from distributions of point estimates derived from bootstrapping. For example, some form a 90% confidence interval by using the value closest to the 5th percentile and the 95th percentile of the parameter values for estimates of the CI endpoints. This is only generally advised when the distribution of point estimates appears symmetric [Efron and Tibshirani, 1993]. Otherwise, more sophisticated approaches are needed. An approach recommended for cases where the distribution of point estimates is asymmetric or where there appears to be potential bias in the point estimate from the original model, an algorithm for bias-corrected and accelerated confidence intervals (BCa) is applied [IBM, 2011]. This confidence interval can be calculated by Matlab [Mathworks, 2013] and SPSS [IBM, 2013].

For the survey analysis results, the standard deviation of the bootstrap interval reveals that the asymptotic model has limited applicability as the confidence intervals estimated by bootstrapping were much wider than the standard error estimates of the parameters from the model building process. An alternative to Wald confidence intervals are approximate confidence intervals based on profile-likelihood estimates centered about Firth's estimate of the model parameters [Firth, 1993]. These profile-likelihood estimates are for many cases more accurate than those derived from traditional Wald confidence intervals [Meeker and Escobar, 1995]. These intervals do not assume asymptotic properties of large samples. These types of confidence intervals are available in standard commercial software [SAS, 2013].

D.9 Individual Factor Analysis Results

This section presents the results of the individual factor analysis. First, results of preliminary single factor tests for significance are presented. Then, for all of the factors that the preliminary tests identify as significant or mildly significant, post hoc comparisons between various sets of factor levels are defined, and the results from Tukey and Scheffé tests for significant differences are discussed. This section also presents results from an analysis of continuous factors.

D.9.1 Results of Preliminary Tests for Categorical Factors

Table D.5 presents the ρ-values that come from analysis of variance, Pearson's Chi-square test and Exact Pearson's Chi-square test.

Table D.5 Statistical Results Overview – Categorical Factors

Factor	ANOVA	Pearson	Exact
Type of Contact	0.086	0.086	0.088
Motive - Profit from Development	0.826	0.822	1.000
Motive - Profit After Development	0.359	0.351	0.410
Motive - Future Work	0.376	0.368	0.545
Motive - Future Relationships	0.851	0.848	1.000
Motive - Satisfy Existing Customer	0.716	0.710	0.784
Motive - Block Competitor	0.327	0.319	0.422
Systems and Software Process Maturity of Respondent Organization	0.501	0.481	0.515
System and Software Process Maturity on Project	0.688	0.668	0.727
System and Software Process Maturity of Prime	0.174	0.169	0.169
Level of Access	0.637	0.620	0.692
Level of Project Manager Experience	0.122	0.122	0.101
Rank - Market Share	0.428	0.409	0.437
Rank - Prestige	0.336	0.321	0.330
Extent of Relationship with Proposal Customer	0.524	0.507	0.520
Level of Customer Satisfaction on Previous or Ongoing Effort	0.037	0.042	0.011

The rows that are greyed with bolded text correspond to the factors for which at least one of the three tests conducted identified a ρ-value less than 0.2. Observe that the factors for which at least one of the tests produced ρ-values less than 0.2 are System and Software Process Maturity of Prime, Level of Project Manager Experience and Level of Customer Satisfaction on Previous or Ongoing Effort.

D.9.2 Post-Hoc Comparisons for Categorical Factors

Based on an analysis of the data, the following contrasts were defined:

1.) Proposals for firm fixed price (FFP) contracts vs. proposals for contracts other than FFP

2.) Proposals where the prime had a System and Software Process Maturity rating of 1 or 2 vs. proposals where the prime had a Systems and Software Process Maturity rating of 3-5

3.) Proposals where the project manager is very experienced vs. proposals where the project manager is less than very experienced

4.) Proposals where the customer is very satisfied with previous or ongoing contract work vs. proposals where the customer is less than very satisfied with previous or ongoing contract work

Contrast #1 above is found to be significant by both Tukey and Scheffé at the 0.05 level. Contrast #4 is found to be significant by both Tukey and Scheffé at the 0.01 level. These contrasts can be interpreted as follows:

1.) Proposals for contracts that are FFP have a lower probability of being awarded contracts than proposals for contracts with arrangements other than FFP.

2.) Proposals submitted to customers that are assessed by the supplier to be very satisfied with previous or ongoing contract work from the supplier are more likely to result in contract awards than proposals submitted to customers who are assessed by the supplier to be less than very satisfied with previous or ongoing contract work from the supplier.

D.9.3 Results for Continuous Factor Analysis

Table D.6 presents estimates for the statistical significance of factors related to the number of contacts between the prime contractor and the customer during the proposal phase. In order to compare proposal efforts for small contracts to proposal efforts for larger contracts, each number of customer contacts is normalized by the total number of engineering hours to be worked on the development portion of the contract.

Table D.6 Statistical Results Overview – Contacts with Customers (Normalized)

Contact Type	ρ-value Bootstrap	ρ-value Firth
Number of written contacts	0.496	0.473
Number of telephone contacts	0.286	0.353
Number of video conference contacts	$0.05 < p < 0.1$	0.341
Number of in-person contacts	0.064	0.289
Total number of contacts	0.328	0.404

Once again, greyed lines with bolded text denote factors that have corresponding ρ-values of 0.2 or less by either bootstrapping or Firth's method. The results indicate the number of video conference contacts and number of in-person contacts are identified as being significant or mildly significant respectively.

Table D.7 presents the results of statistical analysis for the effort spent defining the proposed offering (DPO). The quantity DPORatio is the number of systems engineering labor hours spent defining the proposed offering on the proposal divided by the total number of engineering labor hours to be worked on the development portion of the contract.

Table D.7 Statistical Results Overview – Level of Effort in Proposals (Normalized)

Numerator	ρ-value Bootstrap	ρ-value Firth
DPORatio - all skill levels	0.052	0.060
By Activity (all Skill Levels)		
DPORatio - source requirements	0.271	0.275
DPORatio - system technical requirements	0.045	0.066
DPORatio - architecture	0.044	0.154
DPORatio - cost modeling	0.064	0.167
DPORatio - decision analysis	0.213	0.195
By Skill Level (all Activities)		
DPORatio - beginner skill levels	$0.1 < p < 0.2$	0.173
DPORatio – intermediate skill levels	0.127	0.233
DPORatio - advanced skill levels	0.087	0.144
DPORatio - expert skill levels	0.103	0.199

The results are broken out by activity and skill levels. For example, DPORatio – architecture is the total number of labor hours spent defining and validating the architecture on the proposal divided by the total number of engineering labor hours to be worked on the development portion of the contract. All skill levels are included. Also, DPORatio – intermediate skill level is the total number of labor hours to worked by contributors at an intermediate skill level on defining the proposed offering in the proposal divided by the total number of engineering labor hours to be worked on the development portion of the contract. Observe that DPORatio as well as the DPORatio spent on system technical requirements, architecture, cost estimation and decision analysis are all significant or mildly significant. The DPORatio for all of the skill levels is significant or mildly significant.

D.9.4 Model Diagnostics for Continuous Factor Analysis

For models where the factors are continuous and the outcome is binary, the logistic regression model is a well-accepted and well-understood model to use to make predictions and evaluate hypotheses, etc. Standard textbooks [Kutner et al., 2005] address binary logistic regression models and numerous software packages support binary logistic regression models [Cytel, 2013; Mathworks, 2013, SAS, 2013; IBM, 2013]. The applicability of the logistic regression model can be questionable for “small” data sets. In logistic regression, “small” can vary from fewer than 30 data points to hundreds of data points, depending on the nature of the data (e.g., frequency of each outcome, distribution of the factors) and how well the model conforms with the underlying assumptions of a logistic regression model.

The main objective of the factor-by-factor analysis is to evaluate whether or not each factor is statistically significant. This equates to a hypothesis involving a model parameter or alternatively forming a confidence interval (CI) about a parameter of interest and seeing if the null value (i.e., 0) is included in the CI. Reject H0 only if the null value is not in the CI. The Wald test about a maximum likelihood estimate (MLE) of the parameter of interest [Kutner et al., 2005] provides a reliable and well-accepted way of evaluating such a hypothesis, but only if the use of a logistic regression model is appropriate.

For a logistic regression model to be used, certain assumptions must be met. Table D.8 provides a description of each assumption. For all of the models analyzed, the bootstrapping exercise described above reveals that the data sets were too small to apply the Wald test to determine whether or not the response varies as a function of the covariate. This is the reason why alternative approaches such as bootstrapping and profile likelihood confidence intervals are explored in this analysis for evaluating the significance of candidate factors. Nonetheless, model diagnostics were performed for all of the logistic regression models. In general, other than sample size, most of the diagnostics do not reveal violations of the assumptions for a binary regression model. An example set of these diagnostic plots are provided for illustration for probability of contract award vs. the ratio of systems engineering labor hours spent defining the proposed offering on the proposal to the total number of projected total engineering labor hours on the development portion of the contract (previously labeled DPORatio). The plots presented correspond to the set of responses used in the final effort data quantitative analysis. Preliminary versions of these plots were useful in identifying outliers and influential data points. After careful evaluation, most of these outliers were assessed to be legitimate data points. However, this careful analysis discovered issues with a few data points that eventually led to the exclusion of that data point from the analysis.

Table D.8 Assumptions for a Binary Logistic Regression Model

<table>
<tr><th>Assumption</th><th>How the Assumption was Checked for Each Predictor</th></tr>
<tr><td>Sufficiently large data set</td><td>Bootstrapping was applied for 1,000 runs and the distribution of the key model parameter was compared with that estimated by the logistic regression model for that same model parameter. This approach for evaluating the sufficiency of the data set is recommended by Kutner et al. [2005].</td></tr>
<tr><td>No deviant residuals</td><td rowspan="2">Half-normal probability plot with simulated envelope in which normal values fall within simulated envelope</td></tr>
<tr><td>Adequacy of the linear part of the model</td></tr>
<tr><td>Residuals are well behaved</td><td>Residuals follow two curves and there are no outliers from those two curves</td></tr>
</table>

Figure D.12 displays a histogram of model parameters for the coefficient of DPORatio in the logistic regression model form. In Figure D.12, a coefficient value of zero can be interpreted as the probability of contract award remains constant as the DPORatio increases. A coefficient value of less than zero indicates that the probability of contract award actually decreases as DPORatio increases, and a positive coefficient indicates that the probability of contract award increases as the DPORatio increases. If we can gain statistical confidence that the probability of contract award increases as the DPORatio increases, then we can reject H0. As can be seen, out of 1,000 bootstrap iterations, over 900 of them resulted in an estimate of the coefficient of DPORatio being greater than 0. This is strong indication that the true value is likely greater than zero. Figure D.12 illustrates that the standard deviation of bootstrap estimates is greater than the estimated variation from the asymptotic model. This can be interpreted as the asymptotic model is not appropriate and the estimates of significance from a standard Wald test are not necessarily reliable indicators of the level of significance of the true statistical relationship.

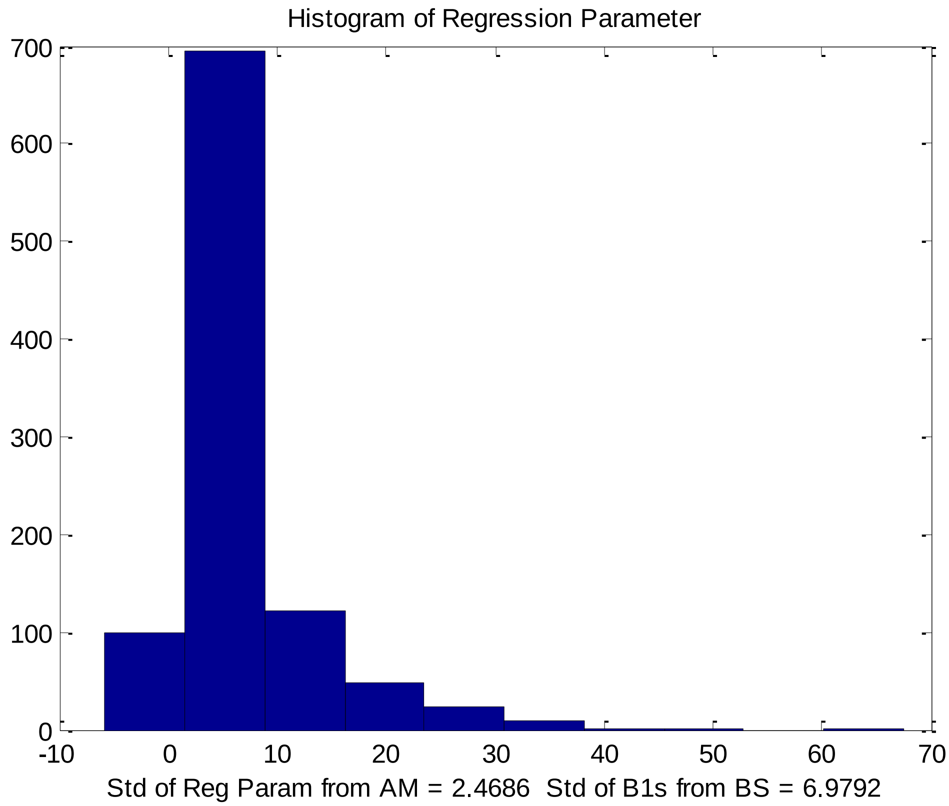


Figure D.12 Example - Histogram of Regression Parameter Derived from Bootstrapping

Figure D.13 is an example of the half-normal probability plot with a simulated envelope. This envelope is simulated using a standard process defined by Kutner et al. [2005]. A half-normal probability plot provides information about whether the normalcy model assumption for a binary logistic regression model holds. Generally speaking, normal values (i.e., stars) that lie below the upper dotted line and above the lower dotted line indicate that normality is O.K.

Figure D.14 depicts a number of different types of residuals for the binary logistic regression model designed to predict probability of contract award based upon DPORatio. These display residuals, Pearson residuals, studentized Pearson residuals, and deviation measurements of the model as defined by Kutner et al. [2005]. For a logistics regression model, the residuals form a distinct pattern. This is because the actual values for each data point are all either zero or one. Therefore, there are generally two curves that form. Residual values that are far away from either of these curves are general indicators of outliers. The curves presented in Figure D.14 generally do not indicate significant issues with the outliers.

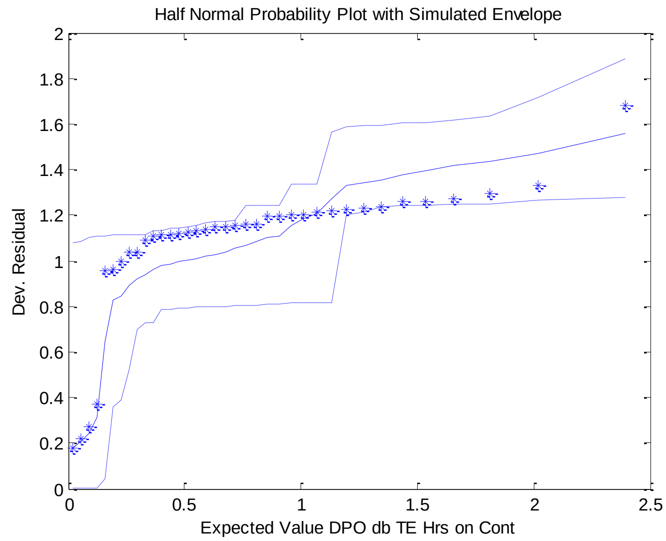


Figure D.13 Example - Half Normal Probability Plot with Simulated Envelope for a Logistic Regression Model

Figure D.15 displays a number of additional diagnostic measurements for a binary logistic regression model. The top left quadrant displays delta Chi-square deviance statistics versus the index of the data point. The top right quadrant displays the delta deviance statistics for the same set of data. For both delta Chi-square and delta deviance, case 33 has an outlying value. This indicates that case 33 is likely highly influential in the model. The lower left hand and right hand quadrants of Figure D.15 plot delta chi-square and delta deviance statistics against the estimated probabilities of contract award. Once again, case 33 (estimated probability of contract award of 0.75) is an outlier. For a thorough explanation of the definition of these diagnostics and a more thorough discussion of how to interpret the values, please see Kutner et al. [2005]. An examination of data point 33 reveals that data point 33 corresponds to a data proposal where a large number of labor hours relative to the contract size were expended. In fact, the binary logistic regression model predicts a 75% probability of contract award based upon this level of effort. However, the proposal effort did not result in a contract award. Because several diagnostic measures identified data point number 33, much scrutiny was put into examining the survey response corresponding to this data point to see if there were any indicators that the respondent did not understand the questions or provided contradictory answers to the questions. This was determined not to be the case. In other words, the author's best judgment is that data point 33 is likely valid. This indicates an important reality of the use of systems engineering in proposal management: if an organization expends a great deal of effort defining the proposed offering, that organization cannot ensure a contract award.

D.10 Multivariate Model to Predict Probability of a Contract Award

This section explores the process of developing a single multivariate binary logistic regression function that will express probability of award as a function of important variables. A multivariate binary logistic function is a function with one dependent variable (i.e., probability of contract award) and potentially multiple independent variables.

The independent variables can be continuous variables that can assume values on an interval or categorical variables that are constrained to assume a finite set of pre-defined values. In the previous sections, factors have been analyzed one at a time for correlation with probability of contract award. Analyzing factors one by one provides a great deal of insight. However, examining factors individually is not adequate to understand the complete set of relationships. In regression analysis, sometimes sets of independent variables will each be highly correlated with some outcome. If the independent variables in the set are highly correlated with each other, it may not make sense to include all of them in a model that can potentially be used to predict the dependent variable. Various well defined and well understood techniques in regression analysis exist to select a subset of independent variables to include in a multivariate regression model.

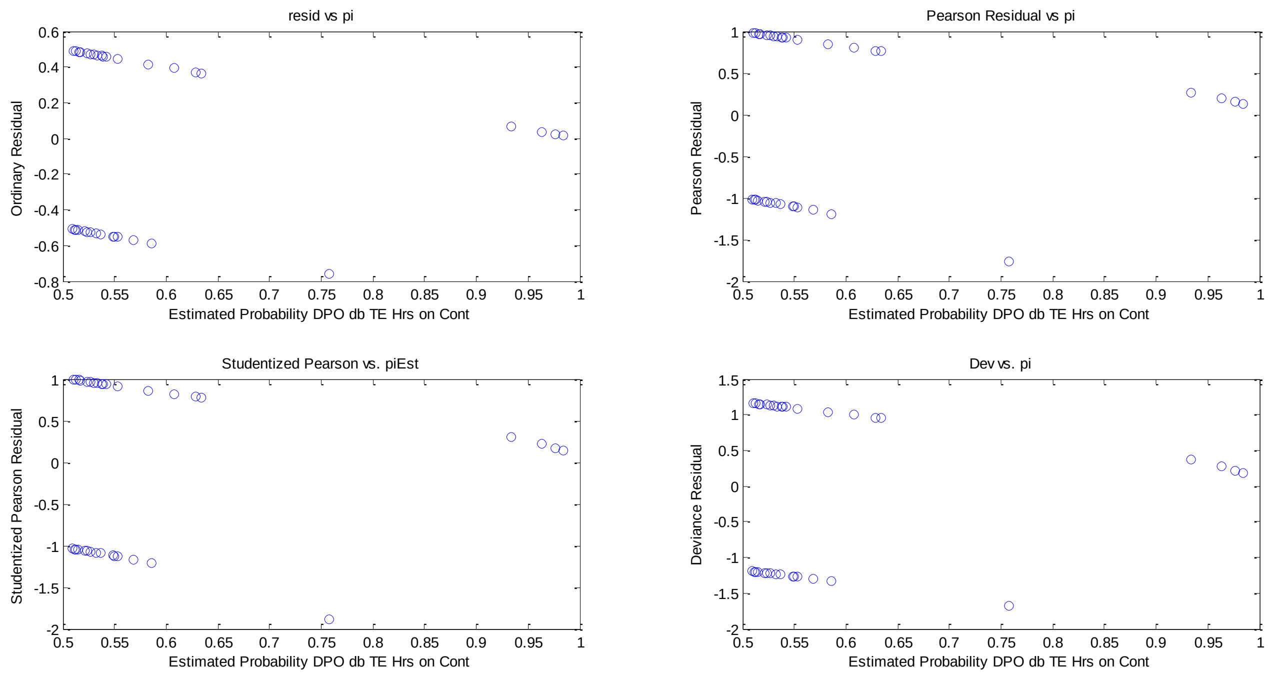


Figure D.14 Example - Plots of Various Residuals for a Logistic Regression Model

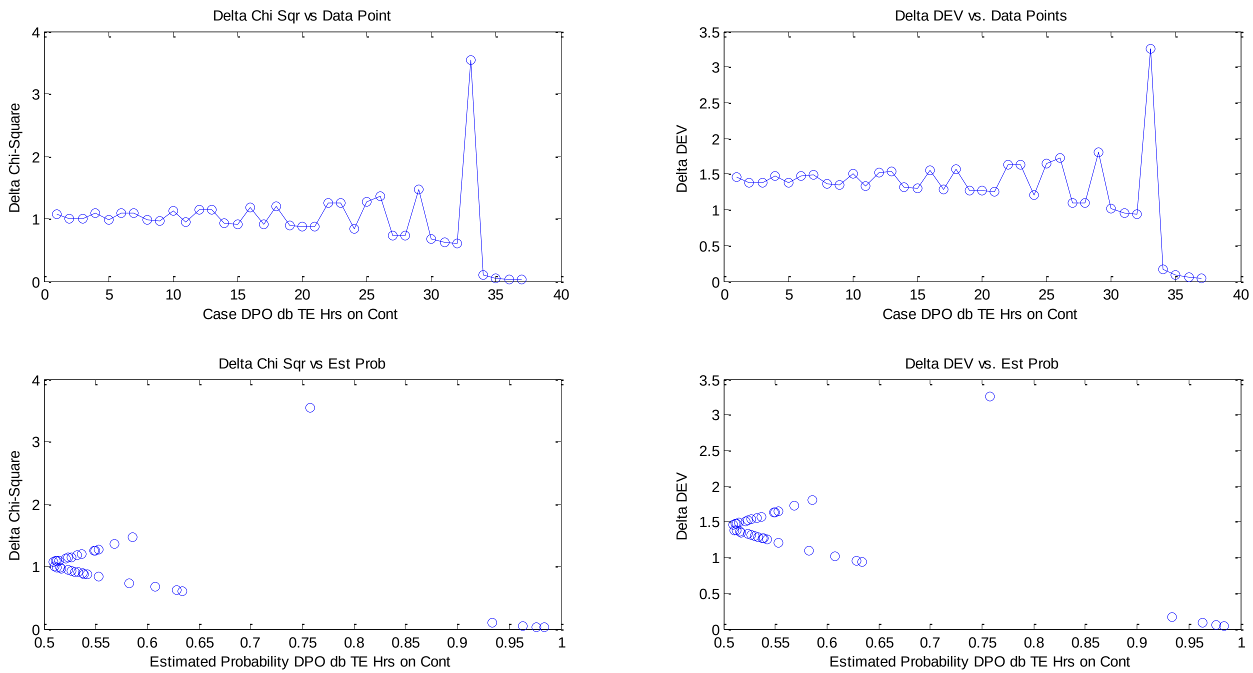


Figure D.15 Additional Diagnostics for a Binary Logistic Regression Model

This section will examine the desired attributes for a multivariate binary logistic regression model. Then, a discussion will follow relating the number of parameters that should be in a binary logistic regression model as a function of the quantity and characteristics of the data used to build the model. The section will then describe the process that is undertaken to develop a binary logistic regression model to estimate probability of contract award as a function of values and levels for a subset of independent variables. This discussion will include diagnostics that are calculated to check various binary logistic regression model assumptions. This section concludes with a discussion of threats to validity related to the model developed.

D.10.1 Goals

A multivariate binary logistic regression model ideally satisfies the following:

1.) The model is "well behaved". As defined here, a well behaved model is a model where the dependent variable is smooth and continuous as the independent variables vary in a continuous way. Since the goal of the DSS is ultimately to find values of the independent variables to optimize the dependent variables, a P function with relatively few local extrema is desirable.

2.) The model accounts for the important factors and interactions between factors that repeat from one data sample to the next. In other words, the model does not under-fit the data.

3.) The model omits factors that are strictly artifacts of the data set used to develop the model. In other words, the model does not over-fit the data.

4.) The model includes parameter estimates that have little or no bias and relatively small variance.

D.10.2 Number of Model Parameters and Data Quantity

The more valid data collected, the more parameters can validly be included in a regression model. A number of researchers address the issue of how many "events per variable (EPV)" are required. An event is defined as the less-frequent (e.g., non-awards) of the two possible binary outcomes [Peduzzi et al., 1996]. This is by definition, less than or equal to half the total number of valid observations. Table D.9 presents an overview of several important findings from the literature addressing the number of parameters that can be used in a multivariate binary logistic model and the number of events in the data.

Table D.9 Minimum Required Number of Events per Variable

Author	Study Findings
Harrell et al., 1985	If there are fewer than 10 EPV, then data reduction methods are necessary. Recommended data reduction methods include: 1. Deleting variables using prior knowledge of the problem 2. Deleting variables known to have high measurement error 3. Principal Component Analysis (if variables are all continuous) 4. Deriving summary indexes after clustering variables either on statistical or subject-matter related grounds
Peduzzi et al., 1989	As EPV < 10, 1. Parameter estimates become biased. 2. Sample variances are both underestimated and overestimated. 3. Derived confidence intervals do not have adequate coverage. 4. In some cases, the signs of the parameter estimates reverse (paradoxical associations). Peduzzi shows that these issues get progressively more severe as the EPV is reduced from 10 to 5 to 2.

When the number of events, which in this case is non-awards, is considered, it quickly becomes clear that the data only supports a model with a limited number of parameters. There are 37 entries assessed to be valid and sufficiently complete. Of those, 22 report contract awards, and 15 report no contract awards. The number of events, therefore, is 15. This means that:

1. For an intercept-only model, the EPV is 15.
2. For a model with an intercept and one additional parameter, the EPV is 7.5.
3. For a model with an intercept and two additional parameters, the EPV is 5.
4. For a model with an intercept and three additional parameters, the EPV is 3.75.
5. For a model with an intercept and four additional parameters, the EPV is 3.

Note that for a categorical factor with *n* levels, the standard practice is to deploy *n*-1 binary variables. For example, for process maturity level, there are five levels, so there would be 4 variables deployed (assuming no additional level is necessary for coding missing data), necessitating four parameters. Including only a subset of these parameters for a single factor is considered poor practice, unless previous analysis has been done that allows aggregation of some of the levels. If process maturity level alone was the only factor considered, the EPV would be 3, as the model would include an intercept and 4 additional parameters.

The same analysis can alternatively provide insight into how many survey responses would be required to develop the ideal DSS. The Ideal DSS includes 20 continuous factors and 7 categorical factors with 4-5 levels each. Main effects only model (i.e., no interaction terms or higher-order polynomial terms) would require approximately 20 (for continuous factors). Assuming each of the 7 factors has an average number of levels of 4.5, including the intercept, this equates to 46 parameters. Assuming data is a 50-50 split between awards and no awards (i.e., best case scenario), to achieve 10 EPV would require 10/0.5*46 = 920 valid data points. If one assumes that the trend of more survey responses report proposal projects that result in contract awards than proposal projects that do not result in contract awards and one assumes that a large percentage of survey responses will be found to either be invalid or insufficiently complete to use for developing a multivariate binary logistic regression model, the actual number of survey responses is far greater than 920. The number of valid survey responses was 62. Many of those 62 responses cannot be used to develop a multivariate binary logistic regression model because they are insufficiently complete. Therefore, it can be concluded that the Ideal DSS cannot be developed from the set of survey responses collected. Therefore, a *P* function must be constructed using a subset of the parameters considered in the Ideal DSS.

D.10.3 Selecting and Evaluating the P Function

This section discusses how to select a subset of parameters to be included in the probability of award function that satisfies the model goals and includes only the essential parameters. This will include a discussion about how to address missing data, a discussion about pre-processing the data, an overview of automated model selection methods and applications of these automated model selection methods to identify promising models. A few key metrics will be defined to aid in selecting a model. Once a model is selected, various diagnostics will be computed to evaluate how well the selected model satisfies the various model assumptions for a multivariate binary logistic regression model. If there are issues with model assumption satisfaction, alternative models will be constructed and evaluated.

D.10.3.1 Addressing Missing Data

One major challenge with using data collected from a survey to develop a multivariate binary logistic regression is that many of the survey responses are missing data. For the survey data, 62 responses were included in the categorical data analysis. Of those 62, 50 are included in the quantitative analysis related to number of customer interactions. Of those 50, 37 are included in the effort data quantitative analysis. There are still some questions unanswered even in those 37 responses. In order to proceed with the analysis, a consistent approach must be adopted for addressing missing data. There are several standard ways statisticians address missing data. These include (1) listwise deletion, (2) imputation, (3) for categorical variables, add a level for non-response.

Listwise deletion means simply deleting all data points that do not have answers for every item. This seems like a safe and conservative thing to do, but has distinct problems. First, listwise deletion does not provide a way to deal with questions that were never asked. The survey included question logic that presented respondents with questions based upon their answers to previous questions. For example, respondents that indicated that there was no previous or ongoing contract work between the prime organization and the customer organization were not asked to rate their perceived level of customer satisfaction for that existing work. In addition, listwise deletion can bias the findings of the analysis by introducing non-response bias. In some studies, not answering a key question has been shown to be more strongly correlated with a given response than selecting any of the provided answer options [Howell, 2013; Sethi and Seligman, 1993].

Imputation is a technique that is used to fill in missing values for statistical analysis. Multiple approaches exist [Schafer, 2005]. These include hot deck imputation and cold deck imputation. In hot deck imputation, values are selected for missing items by randomly selecting a value for the item from the same data set. In cold deck imputation, values are selected for missing items by randomly selecting a value for the item from another data set. Other approaches used for imputation include inserting a sample mean value for wherever the value is missing and selecting values from a custom distribution. A custom distribution reflects the distribution of example data and relationships within the distribution of missing data.

When imputing data, there are other choices that must be made such as whether to perform single imputation or multiple imputation [Schafer, 2005]. Single imputation involves taking one realization per missing value and proceeding with the analysis. Multiple imputation involves repeating single imputation analysis *n* times. For each of *n* iterations, draw one realization per missing value and then take medians or means of model parameters.

In general, imputation has distinct advantages. Imputation mitigates against non-response bias [Schafer, 2005]. Imputation is a mature and well-accepted approach for addressing missing data [Rubin, 1976]. Multiple imputation is generally preferred unless the ratio of missing values to populated values is very low and the sample data set very large [Rubin, 1976].

Imputation also has disadvantages. Often imputation can be perceived as making up data by communities without extensive background in advanced statistical methods [Schafer, 2005]. Also, the success of imputation depends on problem understanding: “naïve imputations can be worse than doing nothing” [Little, 1988; 288].

A third option for addressing missing categorical data is to assign a particular code for missing data and proceed with the analysis. Assigning such a code allows non-response to be formally dealt with and mitigates the risk of the data becoming biased in favor of respondents. The disadvantage to assigning a code for missing data is that each additional level for a categorical variable corresponds to an additional parameter that must be in the model. In order to maintain a fixed event per variable ratio, this requires additional data points. Alternatively, with an already determined number of data points, adding another level ultimately can reduce the number of variables that can be included in a model.

After considering the options, the recommended approach for this research is to listwise delete data points that are missing values for continuous variable and add a level for missing data for categorical variables. Imputation is avoided because the relative increase in the effective data size of the survey is modest. Of the 25 survey responses that could be used if imputation was applied, only 4 are non-awards. Therefore, the effective number of events would increase from 15 to 19.

D.10.3.2 Preprocessing Data

There are several preprocessing steps that should occur before the remainder of the analysis. One step that has already been undertaken is to perform a factor-by-factor analysis to identify which variables are particularly important. This is necessary because of the number of variables in this analysis exceeds what the model selection methods in the statistical software can do an exhaustive search on. If there are too many variables in the search space, the model selection software uses non-exhaustive search algorithms. When this is the case, the risk exists that the true combination of variables that are the optimal set will be overlooked for an alternative, sub-optimal set. Because of this, it is useful to select a subset of the potential variables prior to applying the model selection software.

As previously discussed, eliminating variables based upon a single factor does have risk [Harrell et al., 1985]. The potential exists to prematurely disregard variables that have latent explanatory power. To mitigate against the risk of prematurely omitting important factors, several steps are taken. First, statistical significance screening thresholds are set to be very accepting (e.g., alpha = 0.2). Second, the analysis includes factors that subject matter expertise suggests are particularly important but did not show up as significant in the single factor analysis. Third, include factors for which inspection of the results appears to present a clear trend, yet the trend is not quite pronounced enough to be classified as significant in the single factor analysis.

Continuous variables are all scaled to be order 1 magnitude to increase the numerical stability of the statistical estimation processes. As has been discussed, it is clear that the ideal

DSS with the complete set of independent variables cannot be constructed using the survey data. In order to reduce the number of variables, all of the values related to effort defining the proposed offering on the proposal are aggregated. In other words, one total number is considered in the model that is a sum of all five of the activities considered and all four of the employee skill levels considered.

D.10.3.3 Overview of Automated Model Selection Methods

Several automated model selection methods are used to ensure that multiple options have been considered when searching for the best model. Each of these model selection methods is used to nominate a candidate model where the covariates correspond directly to items measured in the survey. Then, all of the best models are compared against several criteria. One criteria is face validity. The model selected must make sense to experts. Otherwise, the model will not enjoy wide acceptance even if it is highly predictive. Second, the model must score well as compared to other alternatives using several key metrics. One metric of interest in EPV. While the standard EPV value of ten that is widely accepted [Peduzzi et al., 1986] may not be achievable with the limited quantity of data available, models with a higher EPV are generally preferable to models with a lower EPV. Two additional metrics are used to compare models: Akaike's Information Criteria (AIC) and Schwarz Criteria (SC) [Kutner et al., 2005]. AIC and SC are two indices that reward high explanatory power in the model but penalize excessive model terms [IDRE, 2013]. Generally speaking, models with lower values of AIC and SC are considered preferable to models with higher values. Both metrics are used extensively in comparing logistic regression models, and the unique role each metric plays is best exemplified in two quotes. One author writes "we can suggest using the AIC-optimal model for prediction, and the SC-optimal model for description and interpretation" [Shtatland, 2001: 224] and another author observes, "When the more complex model is correct, the AIC selects it more often than does the SC, and when the simpler model is correct, the SC selects it more often than does the AIC" [Ludden et al., 1994: 431].

There are several inherent challenges in using automated model selection methods. One major challenge is that statistical analysis software packages used for this research only run these in conjunction with maximum likelihood estimates (MLEs) of parameters. For small sample data, such as the data collected, Firth's method is preferable to MLE for estimating parameters because it introduces less bias [Maiti and Pradham, 2007]. Therefore, there is no guarantee that models identified as "the best" with parameters estimated by MLE are actually the best models when the parameters are estimated by Firth's criteria. In addition, for all automated model selection methods, there is risk that for small data sets, removal or addition of as few as one data point may change the recommended models.

Table D.10 presents a brief overview of available methods, their advantages and disadvantages.

D.10.3.4 Identifying Promising Models

Each of these four methods displayed in Table D.10 (forward stepwise selection, forward selection, backward elimination and best subsets) are used to identify promising models. Forward selection is run specifying the categorical variables as categorical variables. This is done with main effects only and two variable interactions and quadratic terms. Maximum likelihood estimates for binary logistic regression models with categorical variables are prone to data separation. Data separation is a process where the algorithms for computing the model parameter estimates fail to converge and no model parameter estimates are computed. After preliminary experimentation resulted in data separation, an alternative approach was used for the other model selection methods (i.e., forward stepwise regression, backward deletion and best subsets). In these approaches, to avoid data separation, categorical variables are treated as continuous. This has the risk of any of the model selection methods producing an a model that is not a good model. However, when the models are compared, models that should never have been selected will not compare favorably and thus will be eliminated from the model selection process. For forward stepwise and backward deletion, main effects and two variable interactions/quadratic terms are considered. For best subsets, only main effects models are considered.

D.10.3.5 Selecting the Best Model

After these model selection methods are run and models that do not make sense are filtered out, three models remain. Model 1 has the system and software process maturity rating of the prime contractor and DPORatio included. Model 2 has the systems and software process maturity rating of the prime contractor, DPORatio and the level of customer satisfaction on previous or ongoing efforts. Model 3 has the type of contract and DPORatio. Model 2 minimizes both AIC and SC, but the EPV is 1.45. Model 3 has a slightly higher AIC and SC, but an EPV of 3.25, which suggests that Model 3 is a more stable model than Model 2. Therefore, Model 3 is considered the best option among the alternatives. The results for Model 3 are displayed in Figure D.16.

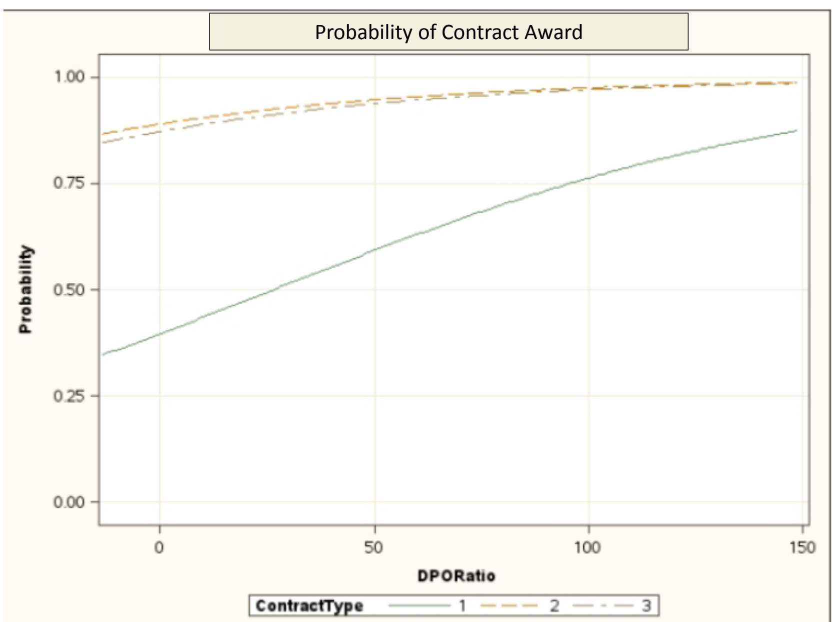


Figure D.16 *P* vs. DPORatio and Contract Type

Table D.10 Overview of Model Selection Techniques

Technique and Definition	Advantages	Disadvantages
Forward stepwise regression – a procedure where covariates are added to a model in descending order of significance and factors are removed that no longer meet the significance criteria to be in the model.	1.) Accommodates continuous or categorical variables. 2.) Does not allow insignificant covariates to remain in the model.	1.) For logistic regression models, data separation is likely. This leads to situations where the parameter estimates are either unavailable or unreliable. 2.) Provides only one model, which may not be the best.
Forward selection – a procedure where covariates are added to a model in descending order of significance until no more covariates meet the criteria for inclusion. This procedure does not remove any covariates that have become insignificant.	1.) Accommodates continuous or categorical variables. 2.) Not likely to result in cases with data separation for logistic regression.	1.) Does not have mechanism for removing covariates that have become insignificant as other terms have been added. 2.) Provides only one model, which may not be the best.
Backward elimination – Start with full model and prune least significant covariates until all the remaining covariates meet the threshold level of significance.	1.) Accommodates continuous or categorical variables. 2.) Does not allow insignificant covariates to remain in the model.	1.) For logistic regression models, data separation is likely. This leads to situations where the parameter estimates are either unavailable or unreliable. 2.) Provides only one model, which may not be the best.
Best subsets – Examine all of the best subsets for models with different numbers of parameters with respect to some goodness of fit criteria.	1.) Provides an array of model options with varying numbers of parameters.	1.) In most packages, the search is only exhaustive with a limited number of variables. 2.) Accommodates only continuous variables, and thus only yields approximate results when used with categorical variables.

In Figure D.16, the curve at the bottom that passes approximately through the point (0,35) and (100,76) corresponds to firm fixed price contracts. The other two curves correspond to cost plus fixed fee and cost plus percentage of costs respectively. The reason the probability of contract award is so high for these curves even when the DPORatio is low is that of the 37 survey responses that were valid and sufficiently complete to include in the multivariate regression analysis, all of the proposal efforts for cost plus fixed fee and cost plus percentage of cost efforts resulted in contract awards. This is not the case for firm fixed price contracts, where 11 of the 24 proposals resulted in contract awards and 13 did not. Because of this, an alternative model is developed where probability of contract award is defined as a function of DPORatio, but only for firm fixed price contracts. This model is displayed in Figure D.17.

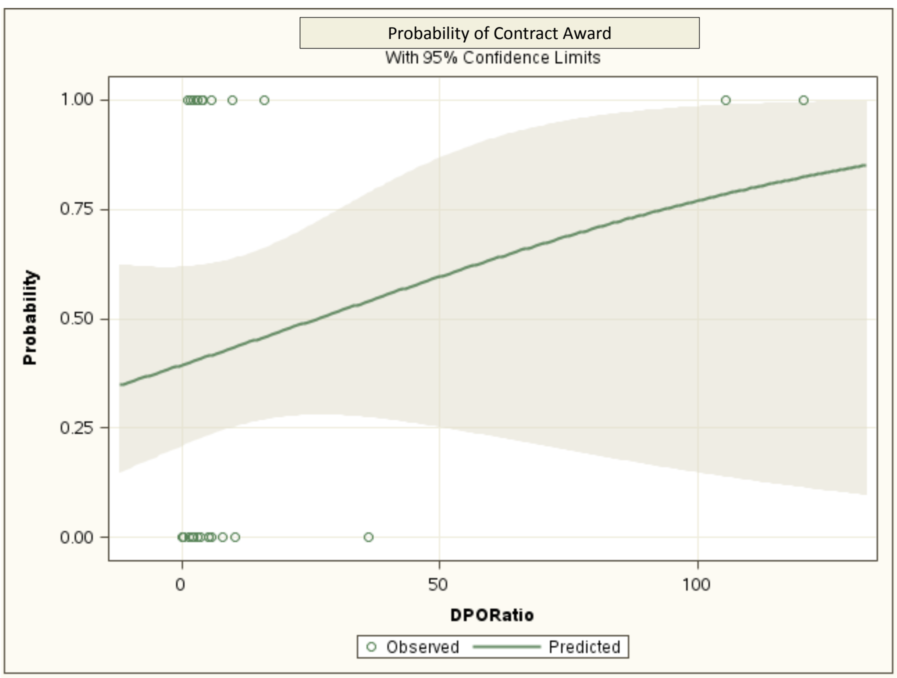


Figure D.17 *P* vs. DPORatio - Firm Fixed Price Contracts

From Figure D.17 it is clear that the probability of contract award increases as the DPORatio increases. Inspection of Figure D.17 suggests that a few outliers may be having a significant influence on the model. Because of this, additional diagnostics need to be computed to understand which of these points are outliers and how influential these outliers are in the model. The following section explores these model diagnostics.

While the trend displayed in Figure D.17 is clearly for *P* to increase as DPORatio increases, the sample size is relatively small, and it is therefore useful to get an idea of the level of statistical confidence that this relationship is indeed increasing. To obtain such an estimate of statistical significance for the relationship, profile likelihood confidence intervals and bootstrapping are used. According to profile likelihood confidence intervals centered about Firth's estimates for the parameters, there is between a 90% and 92.5% chance that the regression is significant and the probability of contract award increases as DPORatio increases. Figure D.18 displays a histogram of bootstrapped estimates for the non-intercept parameter in the *P* vs. DPORatio for Firm Fixed Price contracts.

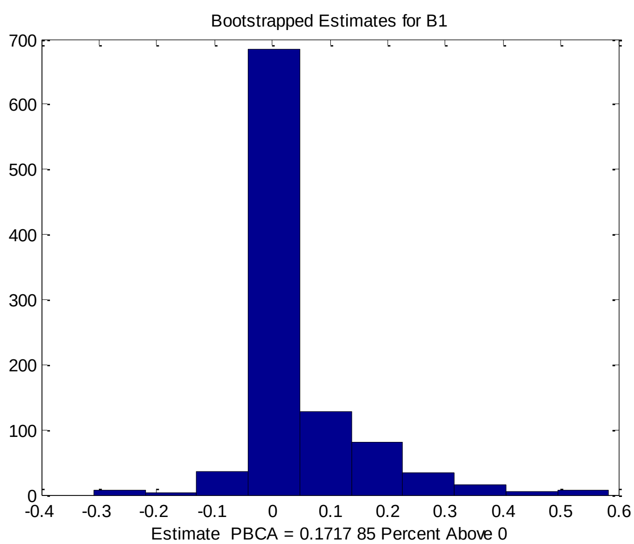


Figure D.18 Bootstrap Estimates of Parameter Values - *P* vs. DPORatio - Firm Fixed Price Contracts

Of the bootstrap estimates of the parameter, 85% are positive and the level of significance derived from bias-corrected and accelerated confidence intervals is 17.1%. Depending on the measure, there is between an 83% and 93% chance that as DPORatio increases, so does *P*.

D.10.3.5.1 Model Assumption Verification Diagnostics for Preferred Model

One standard statistical tool for identifying potential outliers in a distribution is a boxplot. A boxplot is a tool that distinguishes between outliers and non-outliers. A boxplot does not assume any information about the distribution of the data, but rather determines the boundaries between outliers and non-outliers (called whiskers) based on the distance between the quartiles in the data. Figure D.19 displays a boxplot of the DPORatio values presented in Figure D.17.

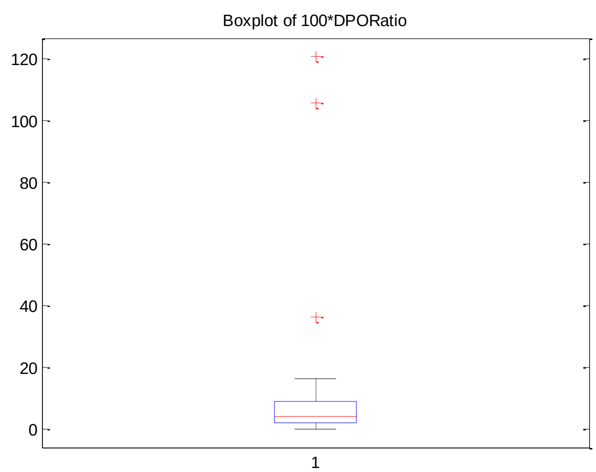


Figure D.19 Boxplot - *P* vs. DPORatio - Firm Fixed Price Contracts

It is clear that three points are identified as outliers. These are the points with DPORatio of approximately 35, 105 and 120. In order to determine if these potential outliers are having an undue influence on the model a couple of items are examined: residuals and leverage values. A plot of the residuals is displayed in Figure D.20. Figure

D.20 displays Pearson residuals, deviance residuals, studentized Pearson residuals and studentized deviance residuals. Detailed definitions of these residuals can be found in the documentation for the SAS tool [SAS, 2013]. Observe that there appear to be no outliers in these values. Figure D.21 displays the leverage values from the model. Two leverage values appear significantly larger than the others.

Observe that there are two extreme values. These two highly influential points correspond to the two most outlying values for which the DPORatio is 105 and 120 respectively. Because of the fact that the most extreme two values appear to have very high leverage in the model and because the boxplot identified the three largest values as outliers, an alternative model is developed where the outliers are not used to develop the model.

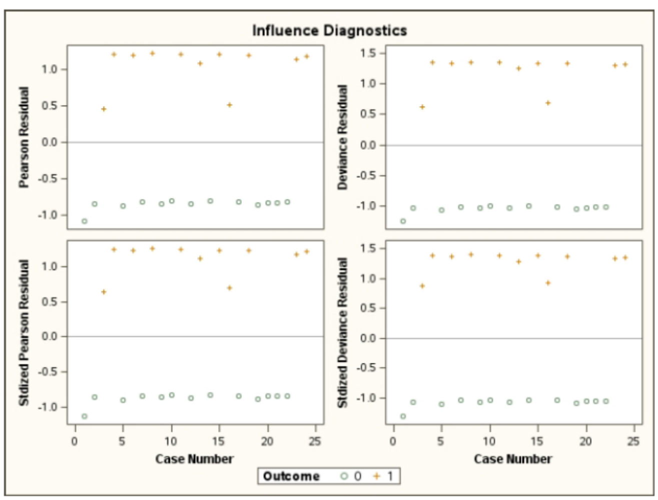


Figure D.20 Residuals - *P* vs. DPORatio - Firm Fixed Price Contracts

D.10.3.6 Model Without Outliers

Figure D.22 displays *P* vs. DPORatio where the three outliers identified by the boxplot analysis in Figure D.19 are removed.

Based upon an analysis of the three outliers there are no apparent errors or anomalies that justify the removal of the outliers. Therefore, the reason that the alternative model without outliers is developed is to better understand the central tendency in the model.

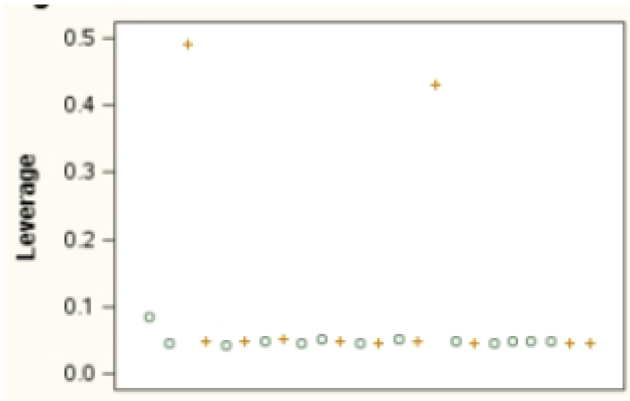


Figure D.21 Leverage - *P* vs. DPORatio - Firm Fixed Price Contracts

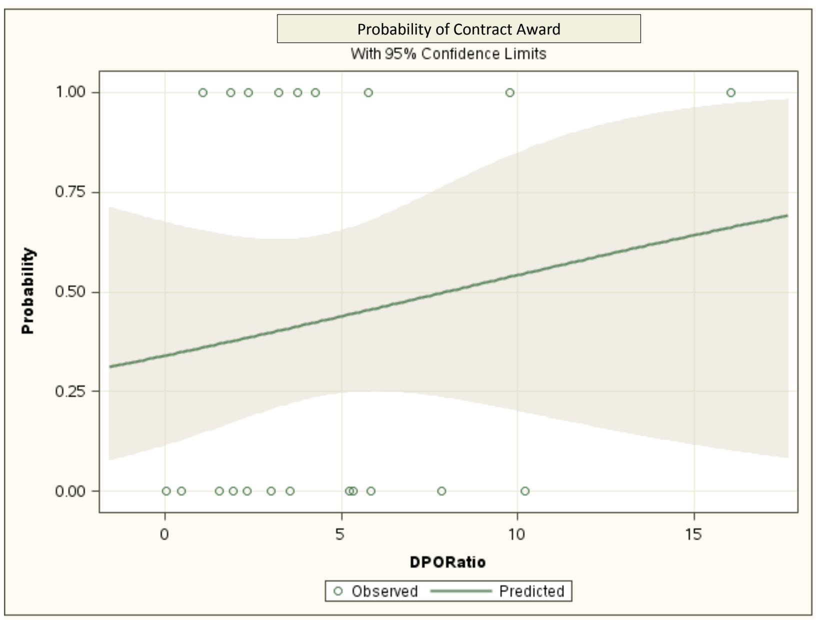


Figure D.22 *P* vs. DPORatio - Firm Fixed Price Contracts - Outliers Removed

As for the model of *P* vs. DPORatio that does not include the outliers, it is necessary to get an estimate of the level of statistical confidence that one can attain that the *P* increases as the DPORatio increases. To obtain such an estimate of statistical significance for the relationship, profile likelihood confidence intervals and bootstrapping is used. According to profile likelihood confidence intervals centered about Firth's estimates for the parameters, there is between a 77.5% and 78% chance that the relationship is positive. Figure D.23 displays the bootstrap estimates for the parameter values.

In Figure D.23 77.6% of the bootstrap estimates are above zero and the estimated level of significance as estimated by bias-corrected accelerated confidence intervals is approximately 25%. This means that depending upon what measure is used, there is between a 75% and 80% chance that the relationship is significant.

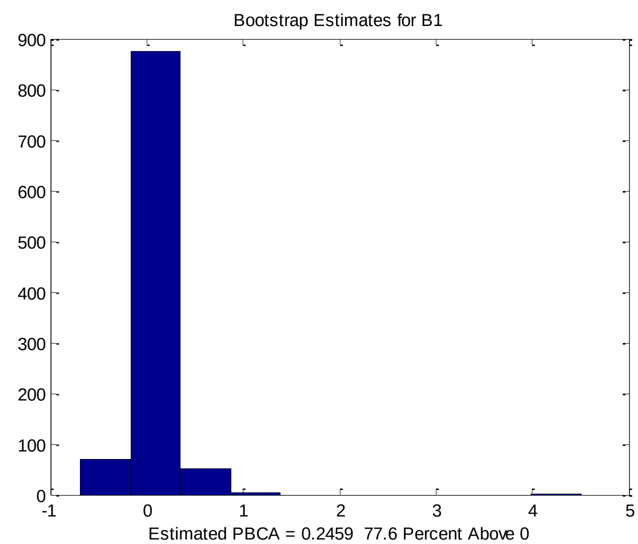


Figure D.23 Bootstrap Estimates – *P* vs. DPORatio - Firm Fixed Price Contracts - Outliers Removed

Appendix E

SEPOMF Detailed Definition and Overview

This section describes the Systems Engineering Proposal Management Optimization Framework (SEPOMF). This includes the type of guidance the SEPOMF provides to decision makers, a definition of the SEPOMF, the SEPOMF objectives, goals and concepts. This section also includes a detailed discussion of the process to apply the SEPOMF.

E.1 Guidance the SEPOMF Offers Decision Makers

The SEPOMF offers prescriptive guidance for decision makers seeking to optimize how much budget to allocate to certain systems engineering activities that are part of defining the proposed offering and what skill level of people need to perform those activities. The SEPOMF can be used to focus exclusively on activities, aggregating all skill levels or focus exclusively on skill levels, aggregating all activities. The SEPOMF provides guidance that takes into account real world constraints on proposal efforts such as limited availability of personnel with particular skills at specific skill levels and overall limitations in budget.

The types of decision makers that the SEPOMF serves are decision makers who must make personnel assignment and resource allocation decisions related to systems engineering on proposals. These decision makers include project manager, functional manager of an organization responsible for systems engineering, lead engineer and lead systems engineer. The definition of these roles as relates to the SEPOMF is described in Table 3.4.

E.2 The SEPOMF Definition, Objectives, Goals and Concepts

The SEPOMF is a modeling framework that helps guide the development of a DSS based upon historical proposal data. The primary objective of the SEPOMF is to allow an organization that is supplying complex systems to maximize the probability of contract award for a future proposal effort by leveraging data related to the use of systems engineering from past proposal efforts.

The goals of the SEPOMF are:

- To provide a concept that is flexible enough that organizations can customize models to use the data that they have, and
- To provide sufficient instructions (e.g., directions, examples) for organizations to implement the SEPOMF.

The SEPOMF is designed to help an analyst create and solve an optimization problem related to the use of systems engineering in proposal management. The SEPOMF uses regression analysis to construct the objective function based upon historical data.

E.3 The SEPOMF Process

While the SEPOMF is a highly flexible and potentially powerful concept, there are many details that must be addressed to effectively apply the SEPOMF. Figure E.1 provides a flowchart that describes the sequence of steps involved in applying the SEPOMF. The steps involved in applying the SEPOMF include identifying decision variables, qualifying historical data, normalizing data, deriving candidate P functions, evaluating candidate P functions, selecting the best P function, defining an optimization problem using the preferred P function, solving the optimization problem, interpreting the results and acting on the model results.

E.3.1 Identifying Decision Variables

A decision variable is defined as “a quantity that a decision maker controls” [OptTek, 2013]. Decision variables differ from other important variables such as the type of contract arrangement or the level of customer satisfaction on recent or ongoing contract work. These other important variables may impact the probability of contract award, but are not under the direct control of the decision maker. For the SEPOMF, the decision variable is an array of number of labor hours to be invested in a set of categories (e.g., activities, skill levels).

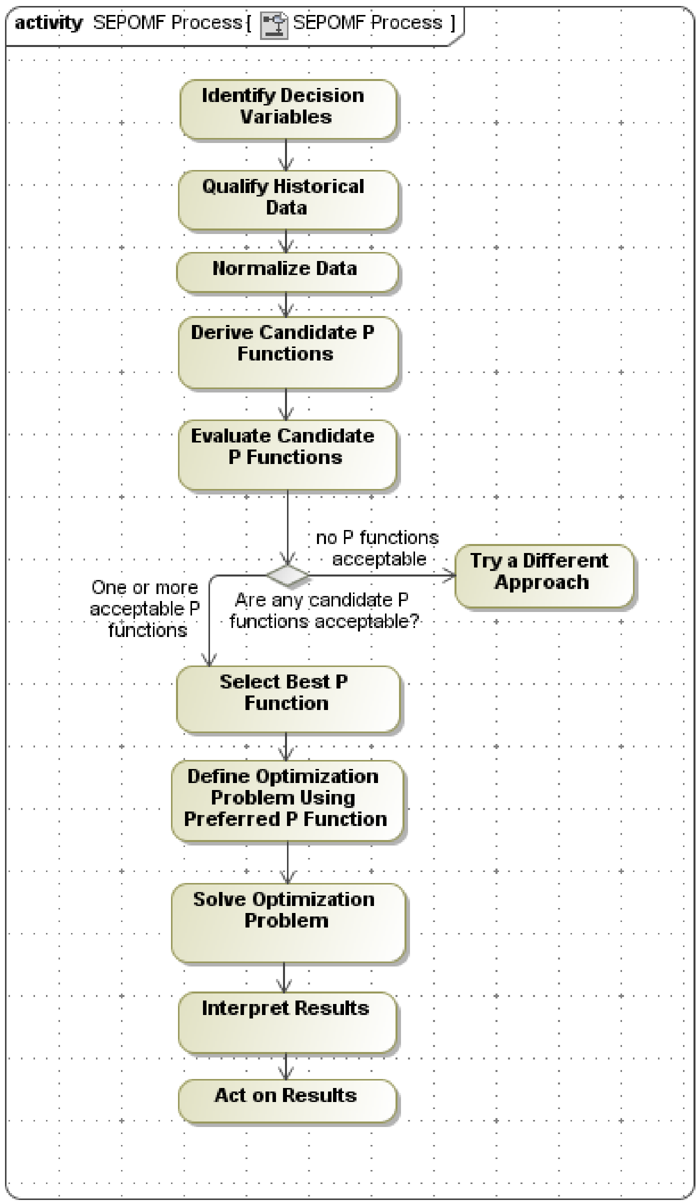


Figure E.1 Process for Applying the SEPOMF

The elements in the array of decision variables should:

- All relate to well-defined items in the problem space
- Be mutually exclusive
- Map clearly to historical data to be leveraged

The elements can be a set of systems engineering activities, a set of skill levels or a combination of systems engineering activities and skill levels.

E.3.2 Qualifying Historical Data

In the SEPOMF, the objective function is derived from regression analysis on historical data. It is important to determine exactly what subset of historical data will be considered in the data set for the regression analysis. This is one of the most challenging judgment calls to make related to the SEPOMF. Some considerations for each possible historical proposal effort are:

1.) Is the historical proposal effort recent enough that similar organizational processes were being followed? If there have been significant changes in processes since the historical proposal effort, the historical data may lead to invalid estimates for parameter values for key relationships in the SEPOMF.

2.) Is the historical proposal effort for the same type of effort as the proposal of interest (e.g., development of a system, maintenance of a system)? If the historical data is for a different type of effort, the historical data may lead to invalid estimates for parameter values for key relationships in the SEPOMF.

3.) Is the historical proposal effort for a similar type of customer to the proposal of interest? Many times different customers follow different processes for acquiring complex systems. If the processes that the customer followed when the historical data was collected are significantly different than what the current customer is following, the historical data may lead to invalid estimates for parameter values for key relationships in the SEPOMF.

E.4 Normalizing Data

It is expected that the historical data will vary (perhaps by orders of magnitude) in terms of contract size and the number of labor hours to be invested on the proposal effort. To effectively leverage historical data with this variance all of the quantitative variables in the optimization model should be ratios of some quantity of interest to some measure of contract size. In the example models, the denominator is the total number of engineering labor hours expected to be worked on the development portion of the contract. Depending upon the nature of the work, some other size factor may be more appropriate. In some cases, the majority of the engineering labor hours are paid for during the proposal and the costs of executing the contract may be primarily driven by production costs. If an organization primarily works these types of contracts, an alternative metric related to number of units sold or total revenue from the contract may be a more useful normalization factor than the number of engineering labor hours for the development portion of the contract.

E.4.1 Deriving Candidate P Functions

The result of the optimization problem will depend upon the objective function. The objective function should include the important variables in the problem and the key relationships between those variables. However, the objective function should not include more variables than the data can support. Considering that the dependent variable in the objective function is probability of contract award (which varies between 0 and 1), a multivariate logistic regression model is a sensible way to analyze the data. The functional form for a logistic regression equation is expressed in Equation E.1.

$$P(x,\beta,\alpha) = \frac{e^{f(x,\beta,\alpha)}}{1+e^{f(x,\beta,\alpha)}}$$

Equation E.1 General Form for Logistic Regression Model

In Equation E.1, the variables have the following meaning:

x – the decision variable – an array of number of labor hours where each entry in the array corresponds to a particular activity and skill level, the values in this array are what is to be optimized

β – model parameters (array)

α – other variables values that are fixed (array)

f – a function of the decision variables, model parameters and other variables

P – probability of contract award

A significant portion of the challenge for deriving *P* is defining the decision variable *x*. A few prominent considerations are what information does the decision maker care about, what aspects of the problem are potentially under the control of the decision maker, and what is the quantity of historical data available. Decision makers can define any number of activities from 1 to *n* and any number of skill levels from 1 to *m*. However, the minimum number of parameters in the model is a function of *n* times *m*. If the decision maker does not care to consider differences in skill levels and just cares about activities or alternatively just cares about skill level mix, then that reduces the dimensionality of the decision variable and likewise the minimum quantity of data to develop a valid model. Even if very large quantities of relevant, historical proposal efforts are available to use to develop a model, reducing the dimensionality of the decision variable will result in higher levels of confidence for the estimates for parameters in the regression model and ultimately higher levels of confidence in the recommendations provided by the model.

The *P* functions will need to include at least linear terms, an intercept and two-factor interaction terms. Let $p = n*m$, which is the total number of components in the decision variable. For example, for a model with five systems engineering activities and four skill levels, p = 4*5 = 20. So for this, there would be at least 20 linear terms, an intercept term and C(20,2) two-factor cross terms (excluding quadratics). C(20,2) = 20!/(2!*18!) = 20*19/2 = 190. Therefore, such a model requires a minimum 211 parameters. As is seen from the literature, one needs ten awards and ten non-awards per parameter. Assuming the historical data has an equal number of awards and non-awards cases, then there is a requirement for 211*(10+10) = 4220 historical proposal efforts. If there are more historical proposal efforts that report contract awards than that report non-awards or vice versa, then the number of historical data points is greater. If there are other important categorical variables included in the model, additional parameters are necessary. In fact, for a categorical variable with q levels, q-1 parameters must be included in the model. If an additional level is specified for non-response, then there are q parameters required. This further increases the requirement for quantity of historical data.

E.4.2 Selecting the Best *P* Function

It may make sense to generate multiple *P* functions and select the "best" function. These functions may vary in the exact set of decision variables considered or the types of cross term factors or higher order polynomial factors to be included. Examine correlations between factors. It may make sense to perform an alternative approach such as principal component analysis if several important variables are highly correlated.

One consideration when selecting a best model is to check for the statistical significance of the terms in the model. Each factor left in the model should be statistically significant. Otherwise, the term may not belong. Some of the model selection methods, such as forward selection methods, will allow terms to remain in the candidate models that may have passed the threshold of statistical significance to be entered into a model, but are no longer significant after other terms are added to the model.

Both quantitative and qualitative evaluation criteria should be used. There are multiple metrics that are used to evaluate a *P* function. There are standard indices that reward high explanatory power but penalize excessive terms. Examples discussed previously in Appendix D include Akaike's Information Criteria (AIC) and Schwarz Criteria (SC). For binary, multivariate logistic regression models another important type of metric is events per variable (EPV). The higher the EPV, the more stable and reliable the model is. Good models have low AIC and SC values and high EPV values.

The primary qualitative concern is that the *P* function should include variables that appear important. Sometimes when there are highly correlated variables, a candidate *P* function will need to include only one of the variables. Selecting the model that appears to include important variables may be a preferable choice. Some may favor models that include variables that are decision variables. Others may favor models that include variables that are meaningful to subject matter experts related to the topic. In either of these cases, care should be taken not to include terms that are not statistically significant even if common sense is that those terms belong. If a decision variable is not statistically significant, it may need to be reconsidered as this is an indicator that adjusting the value of the decision variable will not necessarily change the desired outcome.

Based upon the evaluation criteria, determine if there is at least one suitable *P* function. If there are no suitable *P* functions, consider additional options for defining the objective function. If there is exactly one suitable *P* function, proceed. If there are more than one suitable *P* functions, use the evaluation criteria to select the best one.

E.4.3 Defining and Solving Optimization Problems Using Preferred P Function

After a preferred *P* function (i.e., objective function) is selected, it is important to define an optimization problem that includes constraints. The general mathematical form for the optimization problem can be found in Section 3.5.3. Of course the results of this optimization problem are dependent upon what *P* function is used. For this type of model, there are a number of commercial software packages (e.g., Matlab) that provide functionality to solve optimization problems. A package that offers global optimization options reduces the likelihood of convergence to local optima. These packages generally execute many optimization attempts, each using a different initial value. These global optimization packages use numerical characteristics of the response surface to judiciously select initial values to increase the chances that the global optima will be found by one or more attempt.

E.4.4 Interpreting Results and Acting on Results

After the global optimization problem has been solved and optimal values have been found for the decision variables in the optimization problem, it is necessary to use the various normalization relationships to find values that are meaningful in the problem decision space. For example, in the SEPOMF, the decision variables are ratios of labor hours to be spent on systems engineering activities and or skill levels to a value that represents contract size (e.g., number of engineering labor hours to be worked on the development portion of the contract). The decision maker is not directly interested in the values for this ratio, but rather interested in the absolute number of labor hours that he or she needs to invest in a particular systems engineering activity and/or skill level.

When these values are found, it is important to consider whether the solution appears realistic and believable. A solution that is not acceptable may be an indicator that the definitions for the constraints need to be reconsidered. If the solution does not appear acceptable, reexamine the details of the problem set up to ensure that the problem definition is correct. Also, ensure that the historical data that is being used to derive the *P* function includes contracts, processes and other attributes that are similar to the proposal effort of interest.

When interpreting results it is important to consider threats to validity. Ideally, the parametric relationships in the DSS are derived from statistical relationships that are inferred from historical data. Therefore, any potential for invalid findings from the data analysis threatens the validity of the DSS recommendations. For example, if there is a Type I error in the data analysis and the null hypothesis that a particular factor has no impact on probability of contract award is incorrectly rejected, then the DSS may be developed to recommend that the user take a course of action that in reality does not increase the probability of a contract award.

Table E.1 displays some potential threats to validity when attempting to apply a DSS developed using the modeling framework to a proposal effort of interest. The table includes a name for each potential threat, whether the potential threat is internal or external (see Section 3.13 for definitions), a description of the threat and recommendations for how to mitigate the threat. It is recommended that analysts and decision makers carefully evaluate these threats before attempting to use a DSS developed using the SEPOMF.

Table E.1 Potential Threats to Validity when Applying a SEPOMF DSS

Name	Internal/ External	Description	Mitigations
Non-Causation	Internal	Relationships in the objective function are not causal but simply express correlations between independent and dependent variables. Optimizing these correlated but not causal variables will not achieve the desired improvement in the dependent variables.	With this type of research, controlled experiments are very difficult to conduct. Therefore, it may never be known whether a factor truly causes the probability of contract award to increase. However, the more data that is collected, the better understood the statistical correlations will be.
Local Maximum	Internal	Optimization process converges to a suboptimal local maxima.	Use global optimization. If global optimization software is not available, run optimization routines multiple times using variations in sets of initial values.
Extrapolation	External	Optimization model causes objective function to extrapolate (be evaluated outside the domain of the historical data used to develop objective function).	Analyze diagnostic statistics to identify instances of extrapolation (which may be hidden), and consider other approaches if extrapolation cannot be avoided.
Context Switching	External	There are systematic differences between the set of proposal efforts used to define the objective function and the particular proposal effort being optimized. These can be differences with respect to one of the factors explored in this dissertation or some other factor that appears relevant.	Attempt to collect a complete set of demographic data related to each historical proposal effort so that it will be easy to discern the similarities and differences of the proposal effort of interest from the set of historical proposal efforts. Do not use the dissimilar efforts.
Non-Linear Scalability	External	The strategy for defining the proposed offering may vary considerably between proposal efforts for small contracts and proposal efforts for large contracts. The notion that proposal efforts can be normalized by size such that large proposal efforts can be compared alongside small proposal efforts may not be valid.	When applying the SEPOMF, caution should be taken when comparing historical proposal efforts that differ by orders of magnitude in effort. Further study should be undertaken to characterize how contract size impacts strategy for defining the proposed offering in the proposal.

Appendix F

SEPOMF Example Models

This chapter presents several example decision support systems for the SEPOMF. Providing examples decision support systems developed using the SEPOMF helps make the SEPOMF more concrete. Examples presented have been formalized in mathematical programming language and solved using specialized optimization software (i.e., Matlab with Global Optimization toolbox). These examples are made up for illustrative purposes. The results should not be used to guide policy decisions. Decision makers and their staff can use these examples to help them apply the SEPOMF concept with their historical data. When implementing the SEPOMF, decision makers should use their internal project data to construct *P* functions using regression analysis.

As previously explained, one challenge for using the SEPOMF is the requirement for quantity of data to develop the *P* functions. Based upon feedback from numerous industry experts, many organizations have a limited number of historical proposal efforts that they can leverage to develop *P* functions. Because of this, it is helpful to present several example decision support systems that include a strict subset of the decision variables previously discussed. By reducing the number of decision variables, the required number of historical proposal efforts required also reduces. By presenting these examples, it may become clear how an organization may use its historical data to optimize how systems engineering is used on proposals.

The first two decision support systems focus exclusively on activities and do not include different skill levels (i.e., there is just one skill level) to keep the number of terms in the objective function relatively small. Such decision support systems are useful when there is little difference in the labor costs of the employees at different skill levels and one is really seeking insight related to how to distribute funds across activities.

The third DSS focuses simply on skill levels and does not include activities (i.e., just one activity). This DSS is useful in scenarios where the management does not partition historical data by activity or management wishes to allow the systems engineers to make decisions related to what activities that they need to focus on based upon the maturity of the source requirements, the level of sophistication of the architecture or constituent technologies and the need for precision for cost estimates. Each of the examples contains linear terms, an intercept term and two factor cross terms. Table F.1 presents an overview of the three examples, including what factors are included, the required number of parameters and the required number of proposal efforts.

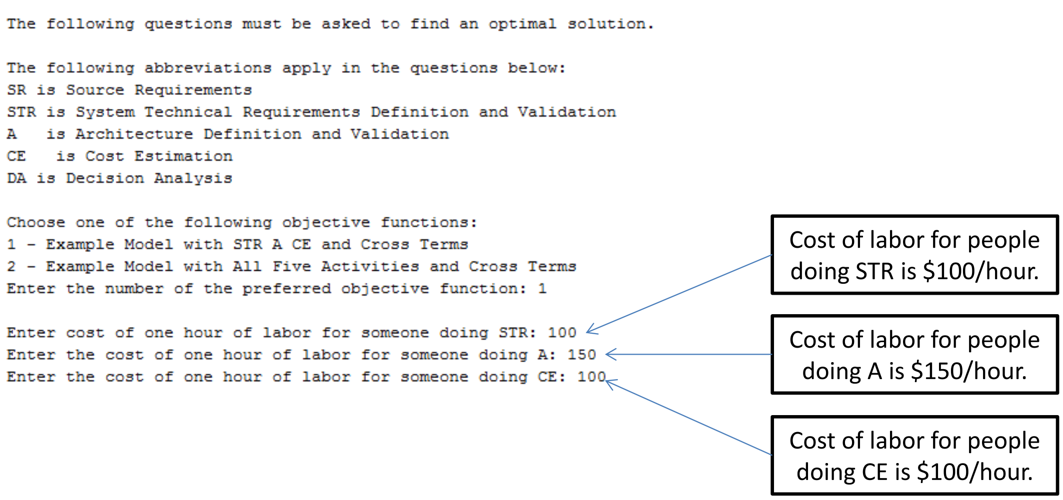
Table F.1 Data Requirements for Example Decision Support Systems

	DSS #1	DSS #2	DSS #3
Factors included	• System Technical Requirements • Architecture • Cost Estimation	• Source Requirements • System Technical Requirements • Architecture • Cost Estimation • Decision Analysis	• Beginner • Intermediate • Advanced • Expert
Number of parameters in objective function	7 terms (1 intercept, 3 linear, 3 two term cross-factors)	16 terms (1 intercept, 5 linear, 10 two term cross-factors)	11 terms (1 intercept, 4 linear, 6 two term cross-factors)
Minimum required number of historical data points to use model	70 proposals where a contract was awarded and 70 proposals where no contract was awarded	160 proposals where a contract was awarded and 160 proposals where no contract was awarded	110 proposals where a contract was awarded and 110 proposals where no contract was awarded

The figures presented in the remainder of the appendix display inputs and outputs from Matlab. In Example DSS #1, the figures include significant detail to explain the various inputs. In Example DSS #2 and Example DSS #3 less explanation is provided as the inputs and outputs can be similarly interpreted.

F.1 Example DSS #1: 3 Activities/1 Skill Level

The first DSS includes three of the five systems engineering activities and aggregates all contributors into one skill level. This may be used by an organization if the historical data just addressed these three activities. Figure F.1 displays the first set of important questions that pertain to how much the hourly labor costs are for contributors who perform each of the activities. In many organizations, hourly labor rates for all of the systems engineering activities are the same. However, this is not always the case. In some cases, the decision maker may have different individuals or groups in mind for each of the activities. Those individuals or groups may differ in their hourly labor rates. Decision support systems developed using the SEPOMF have the flexibility to account for different labor rates for different activities. In this example, the hourly labor rates for system technical requirements and cost estimation is $100/hour and the hourly labor rate for architecture is $150/hour.



```
The following questions must be asked to find an optimal solution.

The following abbreviations apply in the questions below:
SR is Source Requirements
STR is System Technical Requirements Definition and Validation
A   is Architecture Definition and Validation
CE   is Cost Estimation
DA is Decision Analysis

Choose one of the following objective functions:
1 - Example Model with STR A CE and Cross Terms
2 - Example Model with All Five Activities and Cross Terms
Enter the number of the preferred objective function: 1

Enter cost of one hour of labor for someone doing STR: 100
Enter the cost of one hour of labor for someone doing A: 150
Enter the cost of one hour of labor for someone doing CE: 100
```


Figure F.1 Example DSS #1 Inputs (1 of 3)

Figure F.2 displays the next set of important inputs that relate to contract size and budget respectively. As discussed previously, there is a need to normalize different proposal efforts so that proposals for contracts of different sizes can be compared. In these examples, the metric that is used to normalize contracts is number of engineering labor hours to be spent on the development portion of the contract. For this example, the projection is that 10,000 engineering labor hours are to be spent on the development portion of the proposal. The budget is the maximum budget that is to be considered. In this case, the maximum budget is $50,000.

```
The following questions must be asked to find an optimal solution.

The following abbreviations apply in the questions below:
SR is Source Requirements
STR is System Technical Requirements Definition and Validation
A   is Architecture Definition and Validation
CE   is Cost Estimation
DA is Decision Analysis

Choose one of the following objective functions:
1 - Example Model with STR A CE and Cross Terms
2 - Example Model with All Five Activities and Cross Terms
Enter the number of the preferred objective function: 1

Enter cost of one hour of labor for someone doing STR: 100
Enter the cost of one hour of labor for someone doing A: 150
Enter the cost of one hour of labor for someone doing CE: 100
Enter the total number of engineering labor hours planned for the development portion of the contract: 10000
Enter your maximum budget for SE on the proposal (in same units as labor costs): 50000
```

10,000 engineering labor hours are projected on the development portion of the contract

Absolutely no more than $50,000 is to be spent on systems engineering between STR, A and CE

Figure F.2 Example DSS #1 Inputs (2 of 3)

Figure F.3 displays the next set of important inputs that relate to the maximum availability of personnel skilled to perform the various activities. In this example, the least expensive hourly labor rate for any activity is $100/hour. The maximum budget is $50,000. Therefore, it is known a priori that there can be no more than 500 labor hours total worked on the proposal. Therefore, for non-constraining labor availability the value of 500 can be used. In this example, the assumption is that there is definitely enough available talent for system technical requirements and architecture. However, there is only 100 available labor hours for cost estimation.

Figure F.4 displays the first part of the output of the example DSS. In this output, probability of contract award is plotted versus budget in dollars. A note of caution is included in the figure to discourage a reader from attempting to apply these results to real proposal efforts. Once again, this is because the objective function is made up and therefore the output from the DSS is not data-driven. According to this example DSS, if no budget is invested at all, there is approximately a 12% chance that a contract will be awarded. On the other hand, if the entire $50,000 budget is invested, there is a 94% chance that a contract will be awarded. As the budget invested increases, so does the probability of contract award.

There are many available people with STR and A and therefore this is not constraining. Therefore 500 is entered for each.

Only one person is available to do CE and she only has 100 available hours to work to the proposal.

```
The following questions must be asked to find an optimal solution.

The following abbreviations apply in the questions below:
SR is Source Requirements
STR is System Technical Requirements Definition and Validation
A   is Architecture Definition and Validation
CE   is Cost Estimation
DA is Decision Analysis

Choose one of the following objective functions:
1 - Example Model with STR A CE and Cross Terms
2 - Example Model with All Five Activities and Cross Terms
Enter the number of the preferred objective function: 1

Enter cost of one hour of labor for someone doing STR: 100
Enter the cost of one hour of labor for someone doing A: 150
Enter the cost of one hour of labor for someone doing CE: 100
Enter the total number of engineering labor hours planned for the development portion of the contract: 10000
Enter your maximum budget for SE on the proposal (in same units as labor costs): 50000
Enter the maximum number of available labor hours for STR: 500
Enter the maximum number of available labor hours for A: 500
Enter the maximum number of available labor hours for CE: 100
```

Figure F.3 Example DSS #1 Inputs (3 of 3)

Figure F.5 displays the second part of the output. Once again, these results are derived from a made-up objective function and therefore these recommended allocations are made-up. These values or the ratios between these values should not be used to allocate resources on an actual proposal effort. The last few lines of text in the figure provide guidance as to how to distribute the resources. Each column corresponds to a systems engineering activity of interest. Each row corresponds to a particular budget amount. These are the budget amounts that correspond to stars in Figure F.4. To further assist in explaining this output, the row corresponding to a budget of $30,000 is discussed. The constraints that were specified allow all of the budget to be spent. This allows a 55% probability that a contract will be awarded. The recommended allocation of labor hours corresponding to this budget is 149 labor hours on system technical requirements, 34 labor hours on architecture and 100 labor hours on cost estimation. One important point to observe is that the recommended labor allocation honors the constraints that are specified.

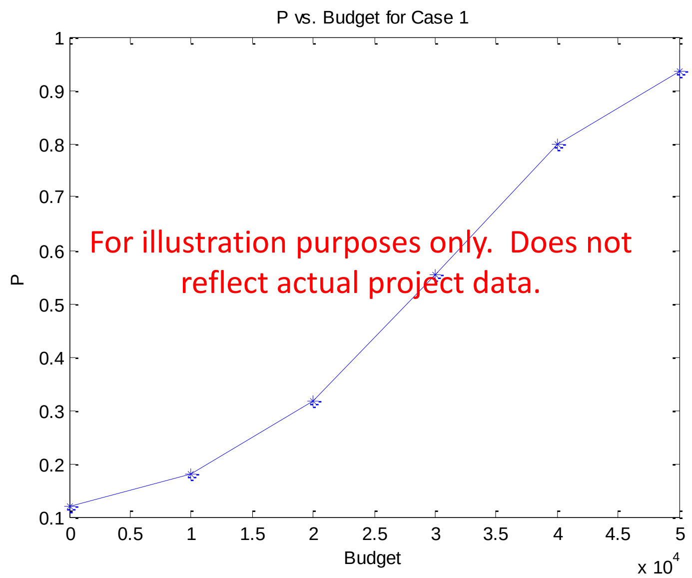


Figure F.4 Example DSS #1 Outputs (1 of 2)

F.2 Example DSS #2: 5 Activities/1 Skill Level

In Example DSS #2 all five systems engineering activities are considered. This example is included primarily to demonstrate that different sets of activities can be considered in decision support systems developed using the SEPOMF. In this example, the assumptions that are made are that all of the activities have the same hourly labor rate and that the labor availability is not a constraint given the budget. Figure F.6 captures the input dialogue for DSS #2.

Figure F.7 displays the probability of contract award versus the allocated budget in dollars. A note of caution is included in the figure to discourage a reader from attempting to apply these results to real proposal efforts. Once again, this is because the objective function is made up and therefore the output from the DSS is not data-driven. For this example, spending no budget once again results in a 12% probability of contract award. Spending the full $50,000 results in an 83% probability of contract award. The more budget that is allocated, the higher the probability of a contract award.

Figure F.8 displays the recommended labor allocation that corresponds to each of the budget levels marked with a star in Figure F.7. As for the previous example, these results are derived from a made-up objective function and therefore these recommended allocations are made-up. These values or the ratios between these values should not be used to allocate resources on an actual proposal effort. The bottom portion of the output describes the recommended labor allocation for each activity. These can be interpreted the same as in the previous example.

```
The following questions must be asked to find an optimal solution.

The following abbreviations apply in the questions below:
SR is Source Requirements
STR is System Technical Requirements Definition and Validation
A   is Architecture Definition and Validation
CE   is Cost Estimation
DA is Decision Analysis

Choose one of the following objective functions:
1 - Example Model with STR A CE and Cross Terms
2 - Example Model with All Five Activities and Cross Terms
Enter the number of the preferred objective function: 1

Enter cost of one hour of labor for someone doing STR: 100
Enter the cost of one hour of labor for someone doing A: 150
Enter the cost of one hour of labor for someone doing CE: 100
Enter the total number of engineering labor hours planned for the development portion of the contract: 10000
Enter your maximum budget for SE on the proposal (in same units as labor costs): 50000
Enter the maximum number of available labor hours for STR: 500
Enter the maximum number of available labor hours for A: 500
Enter the maximum number of available labor hours for CE: 100
Enter the name of the file you wish the model to output to: case1
```

For illustration purposes only. Does not reflect actual project data.

Budget	Spent	P	STR	A	CE
0	0	0.12	0	0	0
10000	10000	0.18	57	0	43
20000	20000	0.32	102	7	88
30000	30000	0.55	149	34	100
40000	40000	0.80	199	67	100
50000	50000	0.94	249	101	100

Recommended labor distribution by level by budget.

Figure F.5 Example DSS #1 Outputs (2 of 2)

```
The following questions must be asked to find an optimal solution.

The following abbreviations apply in the questions below:
SR is Source Requirements
STR is System Technical Requirements Definition and Validation
A   is Architecture Definition and Validation
CE   is Cost Estimation
DA is Decision Analysis

Choose one of the following objective functions:
1 - Example Model with STR A CE and Cross Terms
2 - Example Model with All Five Activities and Cross Terms
Enter the number of the preferred objective function: 2

Enter cost of one hour of labor for someone doing SR: 100
Enter cost of one hour of labor for someone doing STR: 100
Enter the cost of one hour of labor for someone doing A: 100
Enter the cost of one hour of labor for someone doing CE: 100
Enter cost of one hour of labor for someone doing DA: 100
Enter the total number of engineering labor hours planned for the development portion of the contract: 10000
Enter your maximum budget for SE on the proposal (in same units as labor costs): 50000
Enter the maximum number of available labor hours for SR: 500
Enter the maximum number of available labor hours for STR: 500
Enter the maximum number of available labor hours for A: 500
Enter the maximum number of available labor hours for CE: 500
Enter the maximum number of available labor hours for DA: 500
Enter the name of the file you wish the model to output to: example2
```

Figure F.6 Example DSS #2 Inputs

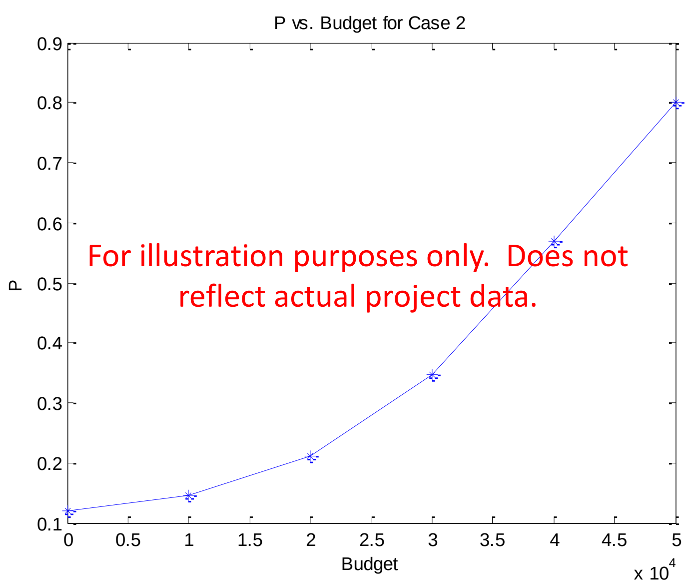


Figure F.7 Example DSS #2 Outputs (1 of 2)

F.3 Example DSS #3: 1 Activity/4 Skill Levels

Example DSS #3 focuses exclusively on skill level mix and assumes just one activity. In many organizations, the historical data records include charges from individual employees but do not designate what activity was being performed that corresponds to each of those charges. Often these can be cross referenced with other data to categorize the employee at a particular skill level, rank or labor grade, depending upon the terminology used in the organization.

```
Choose one of the following objective functions:
1 - Example Model with STR A CE and Cross Terms
2 - Example Model with All Five Activities and Cross Terms
Enter the number of the preferred objective function: 2

Enter cost of one hour of labor for someone doing SR: 100
Enter cost of one hour of labor for someone doing STR: 100
Enter the cost of one hour of labor for someone doing A: 100
Enter the cost of one hour of labor for someone doing CE: 100
Enter cost of one hour of labor for someone doing DA: 100
Enter the total number of engineering labor hours planned for the development portion of the contract: 10000
Enter your maximum budget for SE on the proposal (in same units as labor costs): 50000
Enter the maximum number of available labor hours for SR: 500
Enter the maximum number of available labor hours for STR: 500
Enter the maximum number of available labor hours for A: 500
Enter the maximum number of available labor hours for CE: 500
Enter the maximum number of available labor hours for DA: 500
Enter the name of the file you wish the model to output to: example2
```

For illustration purposes only. Does not reflect actual project data.

Budget	Spent	P	SR	STR	A	CE	DA
0	0	0.12	0	0	0	0	0
10000	10000	0.15	22	30	22	15	10
20000	20000	0.22	42	50	42	35	30
30000	30000	0.36	62	70	62	55	50
40000	40000	0.60	82	90	82	75	70
50000	50000	0.83	102	110	102	95	90

54

Figure F.8 Example DSS #2 Outputs (2 of 2)

Figure F.9 displays the DSS inputs. In this example, the four skill levels described in Table 3.1 are: beginner, intermediate, advanced and expert. In this example, the hourly labor rates increase as the skill level increases. Beginner labor costs \$100/hour, intermediate labor costs \$150/hour, advanced labor costs \$200/hour and expert labor costs \$250/hour. Once again, the assumption is that this is a 10,000 labor hour contract and the maximum budget is \$50,000. As the least expensive labor is \$100/hour, there can be no more than 500 labor hours worked on the proposal. The assumption in this case is that there is adequate available talent that employee availability is not truly a constraint. That is why the value of 500 labor hours is used.

Figure F.10 displays the probability of contract award versus the budget spent in dollars. A note of caution is included in the figure to discourage a reader from attempting to apply these results to real proposal efforts. Once again, this is because the objective function is made up and therefore the output from the DSS is not data-driven. In this example, investing no budget corresponds to a probability of contract award of 12% and investing $50,000 results in a probability of contract award of 62%.

```
The following questions must be asked to find an optimal solution.

The following abbreviations apply in the questions below:
B is Beginner Skill Level
I is for Intermediate Skill Level
A is for Advanced Skill Level
E is for Expert Skill Level n
Enter the cost of one hour of labor for a B: 100
Enter the cost of one hour of labor for an I: 150
Enter the cost of one hour of labor for an A: 200
Enter the cost of one hour of labor for an E: 250
Enter the total number of engineering labor hours planned for the development portion of the contract: 10000
Enter your maximum budget for SE on the proposal (in same units as labor costs): 50000
Enter the maximum number of available labor hours for a B: 500
Enter the maximum number of available labor hours for an I: 500
Enter the maximum number of available labor hours for an A: 500
Enter the maximum number of available labor hours for an E: 500
```

Figure F.9 Example DSS #3 Inputs

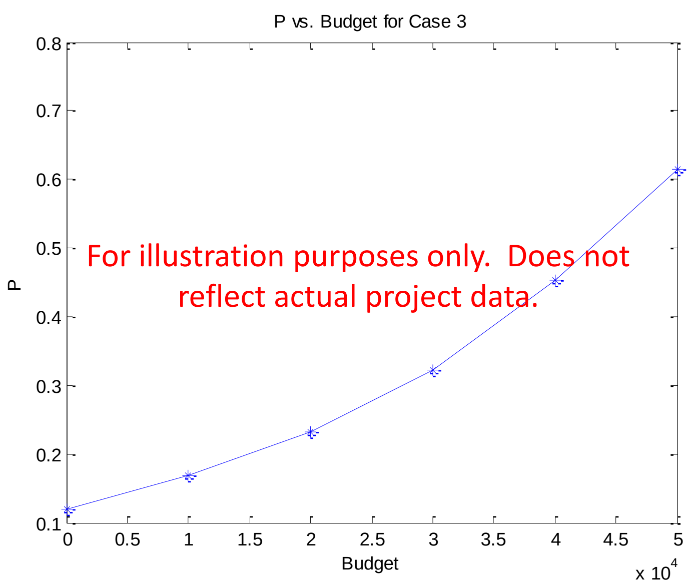


Figure F.10 Example DSS #3 Outputs (1 of 2)

Figure F.11 displays the recommended labor allocation that corresponds to each of the budget levels marked with a star in Figure F.10. As for the previous example, these results are derived from a made-up objective function and therefore these recommended allocations are made-up. These values or the ratios between these values should not be used to allocate resources on an actual proposal effort. The bottom portion of the output describes the recommended labor allocation for each activity. These can be interpreted the same as in the first two examples.

```
The following questions must be asked to find an optimal solution.

The following abbreviations apply in the questions below:
B is Beginner Skill Level
I is for Intermediate Skill Level
A is for Advanced Skill Level
E is for Expert Skill Level n
Enter the cost of one hour of labor for a B: 100
Enter the cost of one hour of labor for an I: 150
Enter the cost of one hour of labor for an A: 200
Enter the cost of one hour of labor for an E: 250
Enter the total number of engineering labor hours planned for the development portion of the contract: 10000
Enter your maximum budget for SE on the proposal (in same units as labor costs): 50000
Enter the maximum number of available labor hours for a B: 500
Enter the maximum number of available labor hours for an I: 500
Enter the maximum number of available labor hours for an A: 500
Enter the maximum number of available labor hours for an E: 500
Enter the name of the file you wish the model to output to: outSkill
```

For illustration purposes only. Does not reflect actual project data.

Budget	Spent	P	B	I	A	E
0	0	0.12	0	0	0	0
10000	10000	0.17	0	0	0	40
20000	20000	0.23	17	-0	-0	73
30000	30000	0.32	67	17	0	83
40000	40000	0.45	114	46	12	77
50000	50000	0.62	161	75	24	72

Figure F.11 Example DSS #3 Outputs (2 of 2)

Appendix G

SEPOMF Validation

This appendix addresses the process of validating the SEPOMF framework. The first part of the appendix will address why it is important to validate a modeling framework and what the process is for validating the framework. Then the subject of exactly what is being validated will be discussed. This will be followed by a description of a process used to derive questions for the validation. Then, demographic information about the validators will be presented as well as a summary of the feedback. The appendix ends with a discussion about what resolution is offered as a result of feedback received in the validation.

G.1 Why Validation is Important

Validation of the SEPOMF, as the validation of most models, is important because it identifies opportunities to improve the SEPOMF's utility and effectiveness by better aligning the framework with reality [Richardson and Pugh, 1981]. Validation can also help uncover incorrect or inapplicable assumptions. Validation of the SEPOMF is not intended to inoculate the model from further criticism. All models, including those developed using the SEPOMF, are by definition approximations of reality and are at best only suitable to support particular types of analysis or decision making.

G.2 Subject of Validation

The subject of this validation is the SEPOMF. The SEPOMF includes the structure of an optimization problem to maximize the probability of contract awards and the definition of systems engineering activities and employee skill levels. Decision makers can invest their budget in employee labor for systems engineering activities on the proposal using employees with varying skill levels. The SEPOMF helps guide the development of optimization models that can be used to determine an optimal allocation of systems engineering labor on a proposal effort from the supplier perspective. The SEPOMF prescribes how to generate an objective function that is based on the historical data available, how to mathematically describe project constraints and how to interpret and solve employee labor allocation optimization problems using specialized mathematical software. Each model developed using the SEPOMF has the potential to help decision makers identify an optimal use of systems engineering on proposal efforts to maximize the probability of being awarded a given contract. More specifically, the solution from each model helps decision makers determine how many labor hours to devote to systems engineering efforts in defining the proposed offering. In some cases, the DSS will also provide recommendations about how to optimally distribute those labor hours across a number of activities and what the optimal mix of skill levels of contributors is for performing each activity.

The subject of this validation is the SEPOMF. While the general concept surrounding the SEPOMF may have application optimizing investments in a portfolio of investment options to maximize the probability of achieving some binary outcome, the focus of this validation is applications of the SEPOMF directly related to the use of systems engineering defining the proposed offering. The decisions that are being explored are how much budget to allocate to systems engineering on proposals and how to distribute that budget.

For the validation, two example models are demonstrated using the SEPOMF. These are DSS #1 and DSS #2 described in Appendix F. DSS #3 was added later to provide an example DSS that focuses on skill levels as opposed to activities. The term “model” is used in this appendix versus DSS because “model” was the terminology used in the validation exercise as opposed to “DSS”.

G.3 Validation Plan

This section describes the required minimum qualifications for validators and provides an overview of the process for validation.

G.3.1 Required Minimum Qualifications for Validators

Because the validation is for a modeling framework or methodology, validators need to understand both the problem domain of the model (i.e., the use of systems engineering on proposals) and the process of developing optimization models where the objective function is derived from statistical regression analysis. Because of this, the following is the set of minimum requirements for someone to be a validator of the SEPOMF:

1. Experience on multiple proposals - Must have contributed to at least five different proposals from the perspective of the supplier in which complex systems are being engineered
2. Systems engineering process expertise - Must have expertise related to systems engineering processes
3. Systems engineering proposal experience - Must have experience using systems engineering on proposals for complex systems
4. Optimization model understanding - Must have an understanding of optimization models
5. Regression model understanding - Must have an understanding of statistical regression models

G.3.2 General Process for Validation

This section describes the validation process of the SEPOMF. First, an email is sent to contacts whose experience suggests that they are likely to meet the minimum requirements for participation in the validation exercise. The email contains a description of the purpose of the validation exercise, the minimum qualifications for validating the model and an estimate of the time commitment to validate the model. If the contacts respond and indicate their interest, then two documents are sent to the potential validator. First, validators are provided a power point presentation about the SEPOMF. Second, validators are provided a validation package. The validation package is a brief document that provides a concise textual description of the research, provides directions and defines a set of questions that the validator will eventually answer. The validator is asked to schedule a telephone conversation with the researcher. One-on-one validation sessions are used for validation to encourage validators to openly ask questions and express opinions. It is expected that in many cases validators may not feel open to communicate freely if they are collocated with individuals employed by other companies (in some cases, potential competitors).

During the telephone conversation, the researcher describes and discusses the presentation. The slides include screen captures of inputs and outputs of example models. The validator is allowed an opportunity to ask the researcher any question he or she wishes to ask. After the telephone conversation, the validator fills out a questionnaire and returns it to the researcher. One question on the questionnaire is whether the validator will entertain follow-up questions. If the validator indicates that follow-up questions are O.K., then the researcher may follow up. If not, the researcher will analyze the input received from the validators as-is.

G.4 Derivation of Questions for Validators

This section explains the process that was used to derive the questions that validators are asked.

G.4.1 Questions about the Qualifications of Respondents

This section discusses the questions that are asked for validators relating to each of the minimum qualifications described previously. Validators are provided a set of minimum qualifications in the initial contact email. It is up to the potential validators to verify that they meet these qualifications prior to agreeing to do the validation. If their responses to these questions indicate that they do not meet the minimum qualifications, then their responses to the questions about the SEPOMF are not used. The qualifications of respondents are reported in Table G.3 through Table G.6.

G.4.1.1 Experience on multiple proposals

One important screening question is asked to verify that validator's experience spans multiple proposal efforts. If the response to this question is a number less than five, the validator inputs are disqualified.

Question for Validators: Approximately how many proposal efforts have you contributed to from the supplier perspective where complex systems are being engineered? If you have been involved in the role of supporting one or more proposal efforts from the acquirer perspective, please do not include those efforts in this estimate.

G.4.1.2 Systems engineering process expertise

Respondents must affirm their understanding of systems engineering processes in order for their inputs to be considered.

Question for Validators: Do you have expertise in systems engineering processes and why do believe this?

G.4.1.3 Systems engineering proposal experience

Respondents must have experience in the use of systems engineering on proposals for complex systems.

Question for Validators: Have you used systems engineering on proposals for complex systems? If so, please provide a brief description.

G.4.1.4 Optimization model understanding

Respondents must affirm their understanding of optimization modeling in order for their inputs to be considered.

Question for Validators: Do you understand the basic concepts of optimization modeling and why do you believe this?

G.4.1.5 Understanding of statistical regression models

Respondents must affirm their understanding of regression models in statistics in order for their inputs to be considered.

Question for Validators: Do you understand the basic concepts of regression modeling in statistics and why do you believe this?

G.4.2 Questions about the SEPOMF

Determining the set of questions to ask validators for evaluating an optimization modeling framework presents a challenge. To begin with, the subject of the validation in this research is a modeling framework versus a single, valid model. As a result, feedback about validator's opinion of the modeling framework is valuable. On the other hand, some detailed questions about the degree to which an attribute of a model effectively represents an attribute of the system may not be relevant in some cases.

No literature was found that provided good questions to ask validators for optimization modeling frameworks or even optimization models. Therefore, literature addressing validation questions from another modeling paradigm, system dynamics models, is used to derive questions. Richardson and Pugh [1981] provide significant insight into useful model tests and associated questions for system dynamics models. Of course, there is not an exact analog for every Richardson and Pugh question to an optimization modeling framework. Table G.1 provides an overview of the Richardson and Pugh validation model tests. After Table G.1 is presented, each of the questions from the Richardson and Pugh model tests is analyzed for applicability to an optimization modeling framework. Where there is applicability, particular questions customized to evaluate the SEPOMF are derived.

Richardson and Pugh classify their model tests into three categories: consistency, suitability and model utility and effectiveness. Merriam-Webster defines the word consistency as "agreement or harmony of parts or features to one another or a whole" [Merriam-Webster, 2013]. Richardson and Pugh equate model consistency with whether or not the model represents the "slice of reality" it is attempting to replicate [Richardson and Pugh, 1981, 312]. While Merriam-Webster does not provide a definition of the word "suitability", they do define the closely related word "suitable" as "adapted to a use or purpose" [Merriam-Webster, 2013]. Richardson and Pugh equate a model's suitability with whether or not a model is suitable for its purposes and the problem it addresses.

Merriam-Webster defines utility as "fitness for some purpose or worth to some end" [Merriam-Webster, 2013]. Merriam-Webster does not define effectiveness but defines the closely related word effective as "producing a decided, decisive, or desired effect " [Merriam-Webster, 2013]. Richardson and Pugh acknowledge that the concept of model utility and effectiveness is very closely related to model consistency and suitability, but assert that model utility and effectiveness goes further. According to Richardson and Pugh, a model with utility and effectiveness communicates to decision makers, generates understanding of the system, enhances insight and influences the audience.

Richardson and Pugh have model tests for both structure and behavior. None of Merriam-Webster's definitions for structure are an exact match for what is being described here. However, the Oxford dictionary defines structure as "the arrangement of and relations between the parts or elements of something complex" [Oxford Dictionary online, 2013]. Behavior is defined by Merriam-Webster as "the way in which something functions or operates" [Merriam-Webster, 2013]. By examining combinations of tests for both structure and behavior in the context of model consistency, model suitability and model utility and effectiveness, Richardson and Pugh are able to define model tests that address a broad array of concerns related to models in general. The next part of this section examines Richardson and Pugh model tests and their applicability to an optimization modeling framework.

In this section, each of the questions in Table G.1 is discussed in the context of optimization models such as those generated using the SEPOMF. The next seventeen subsections examine the applicability of each question derived from Richardson and Pugh in the context of the SEPOMF. The section following the seventeen subsections explores an additional question that will be asked of validators, but does appear to have a direct analog in the Richardson and Pugh model tests.

G.4.2.1 Structure - Dimensional consistency

Question from Richardson and Pugh: Do the dimensions of the variables in every equation of the model agree with the computation?

Analysis: No questions are asked of validators because this should be assessed by the researcher and verified by the dissertation committee and not fielded to validators.

G.4.2.2 Structure - Extreme condition tests in equations

Question from Richardson and Pugh: Does every equation in the model make sense even if subjected to extreme but possible values of its variables?

Table G.1 Richardson and Pugh Model Tests Overview

Activity Type	Test	Relevant Questions
Consistency	Structure - Dimensional consistency	1.) Do the dimensions of the variables in every equation of the model agree with the computation?
	Structure - Extreme condition tests in equations	2.) Does every equation in the model make sense even if subjected to extreme but possible values of its variables?
	Structure - Boundary structural adequacy tests	3.) Does the structure of the model contain variables and feedback effects necessary to address the problem and suit the purposes of the study?
	Behavior – Parameter (in)sensitivity	4.) Is the behavior of the model sensitive to reasonable variations in the parameters?
	Behavior – Structural (in)sensitivity	5.) Is the behavior of the model sensitive to reasonable alternative formulations?
Suitability	Structure - Face validity	6.) Does the model's structure look like the real system? 7.) Is the model a recognizable picture of the real system? 8.) Are those who know the system most closely convinced that a reasonable fit exists between the rate/level/feedback structure of the model and the essential characteristics of the real system?
	Structure - Parameter validity	9.) Are the parameters themselves recognizable in terms of the real system?
		10.) Are the values selected for the parameters consistent with the best information available about the real system?
	Behavior – Replication of reference modes	11.) Does the model endogenously reproduce the various reference modes that initially defined the study?
	Behavior – Surprise behavior	12.) Does the model under some test circumstances produce dramatically unexpected behavior, not observed in the real system?
	Behavior – Extreme conditions	13.) Does the model behave reasonably under extreme conditions or extreme policies, even ones that have never been observed in the real system?
	Behavior – Statistical tests	14.) Does the model output have statistically like data for the real system?
Model Utility and	Structure – Appropriateness of structure	15.) Is the size of the model, its simplicity or complexity, and its level of aggregation or richness of detail appropriate for the audience for the study?

Table G.1—*Continued*

Effectiveness	Behavior – Counterintuitive Behavior	16.) In response to some policies, does the model exhibit behavior that at first contradicts intuitions and later, with the aid of the model, is seen as a clear implication of the structure of the system?
	Behavior – Generation of insights	17.) Is the model capable of generating new insights, or at least the feeling of new insights, about the nature of the problem addressed and the system within which it arises?

Analysis: The objective function in each SEPOMF model is a regression equation that expresses probability of contract award as a function of a number of variables, such as a set of systems engineering activities or combinations of activities and particular skill levels. In general, regression equations are only valid to interpolate but not to extrapolate. In order to interpolate, a regression equation should only be evaluated within the domain of the data that was used to develop the regression equation. The example models use regression parameters that are made up to illustrate the SEPOMF concept. Therefore, they are not derived from any real set of data and there is no defined boundaries for the domain in which the examples are valid. However, the majority of the survey respondents reported less than 1% of the estimated total number of engineering labor hours to be worked on the development portion of the contract for each of the five activities surveyed. Therefore, the parameters were selected to provide useful results for cases where the total number of systems engineering labor hours to be worked on the five activities defining the proposed offering (source requirements definition and validation, system technical requirements definition and validation, architecture definition and validation, cost estimation and decision analysis) total less than 5% of the estimated total number of engineering labor hours to be worked on the development portion of the contract. The example models may produce results that appear unreasonable if used for cases where the ratio of the total number of systems engineering labor hours to be worked on the five activities defining the proposed offering to the estimated total number of engineering labor hours to be worked on the development portion of the contract exceeds 5%.

No questions are asked of validators related to the example models because the focus of the research is on the framework, not on example models made up to help explain the framework.

G.4.2.3 Structure - Boundary adequacy

Question from Richardson and Pugh: Does the structure of the model contain variables and feedback effects necessary to address the problem and suit the purposes of the study?

Analysis: An optimization model such as the SEPOMF is static and does not include feedback effects. However, the optimization model should include the important variables in the problem. The important variables need to come as close as possible to exhaustively spanning the decision space and be mutually exclusive. In other words, there should be no overlap in the definition of the decision variables. Therefore, it is useful to ask validators whether the correct set of variables is included in the model. The variables that are used in the SEPOMF formulation are the number of labor hours to invest in combinations of activities and skill levels. The activities are define and validate source requirements, define and validate system technical requirements, define and validate architecture, cost estimation and decision analysis. The skill levels are beginner, intermediate, advanced and expert. The most expedient way to do this is to ask if there are important variables missing or if there are variables in the model that should not be included.

Question for Validators: Are there important variables that are missing in the SEPOMF formulation that should be included? If so, please explain.

Question for Validators: Are there variables included in the SEPOMF formulation that are extraneous and should be removed? If so, please specify which variables and explain why you believe that they are extraneous.

G.4.2.4 Behavior – Parameter (in)sensitivity

Question from Richardson and Pugh: Is the behavior of the model sensitive to reasonable variations in the parameters?

Analysis: As the SEPOMF models are optimization models and do not have dynamic behavior, this question is not really applicable. The closest potential questions to ask relate to whether outputs of the model vary appropriately as a function of inputs or relate to whether changes in input values generate reasonable changes in output values. However, the example optimization models feature objective functions that have been made up to illustrate the concept. While care was taken to select parameter values that the author believes produce sensible looking recommendations, by definition, these example models are made up.

No questions are asked of validators because validators' opinions about whether the relationships between inputs and outputs for the made up example models make sense have limited value.

G.4.2.5 Behavior – Structural (in)sensitivity

Question from Richardson and Pugh: Is the behavior of the model sensitive to reasonable alternative formulations?

Analysis: Once again, models developed using the SEPOMF are static optimization models and do not have behavior per se.

No questions are asked of validators because the question is inapplicable.

G.4.2.6 Structure - Face validity (1 of 3)

Question from Richardson and Pugh: Does the model's structure look like the real system?

Analysis: Structure for an optimization model has a slightly different context than for a systems dynamics model. In an optimization model, the structure pertains to the relationships between variables in the objective function and constraints. This is specific to a model versus a general modeling methodology. The two models feature objective functions that are made up to illustrate the concept, and need not be validated because they are made up. However, the general forms of the equations for the models presented in the mathematical programming formulation of the problem should make sense to a validator. Therefore, this should be assessed in the questionnaire.

Question for Validators: Do the equations and inequalities (e.g., the constraints) presented in the mathematical programming formulation of the optimization model appear to reasonably represent the variables and relationships related to the use of systems engineering on proposals? If not, please describe issues that you see.

G.4.2.7 Structure – Face validity (2 of 3)

Question from Richardson and Pugh: Is the model a recognizable picture of the real system?

Analysis: In an optimization model, the degree to which the model approximates the real system is best assessed by whether the right set of variables are included in the model and whether the right relationships (i.e., equations and inequalities) are included in the model. Both of these have been asked in previous questions.

No additional questions are asked of validators related to this question.

G.4.2.8 Structure – Face validity (3 of 3)

Question from Richardson and Pugh: Are those who know the system most closely convinced that a reasonable fit exists between the rate/level/feedback structure of the model and the essential characteristics of the real system?

Analysis: Obtaining an answer to this question is clearly an important goal of the validation exercise. However, as there are no rate/level/feedback constructs in optimization models, the question needs to be rephrased.

Question for Validators: Do you believe that a reasonable fit exists between the variables and structure of the SEPOMF and the essential characteristics of the use of systems engineering on proposals?

G.4.2.9 Structure - Parameter validity (1 of 2)

Question from Richardson and Pugh: Are the parameters themselves recognizable in terms of the real system?

Analysis: This is a question that should be applicable to almost any model, including models developed using the SEPOMF. This question should be asked related to the variables in the formal mathematical programming specification of the general optimization model formulation in the SEPOMF.

Question for Validators: Are the variables presented in the mathematical programming formulation of the optimization model recognizable in terms of the use of systems engineering on proposals? If not, please explain.

G.4.2.10 Structure - Parameter validity (2 of 2)

Question from Richardson and Pugh: Are the values selected for the parameters consistent with the best information available about the real system?

Analysis: Parameter values are specific to particular models. The particular example models that the validators are shown feature objective functions where both the relationships between the variables and the parameter values are made up to illustrate the SEPOMF concept. These parameter values have been selected to achieve model outputs that are consistent with the author's understanding of the role of systems engineering in proposal management and are chosen to produce believable results. However, because the parameter values are made up, it is not valuable to know whether validators agree with the author's perception.

No questions are asked of validators.

G.4.2.11 Behavior – Replication of reference modes

Question from Richardson and Pugh: Does the model endogenously reproduce the various reference modes that initially defined the study?

Analysis: In general, the model developer has record of the reference modes and is probably the best positioned to evaluate how closely a model is able to replicate those reference modes. The SEPOMF, however, generates static optimization models, and static optimization models do not have reference modes. There is some analogy for the objective functions used in the SEPOMF models fitting real data. Ideally, these objective functions are regression models derived from data. Therefore, examining residuals or comparing the objective function's estimate of the probability of a contract award to the real award status can gauge how good of a fit the regression model is to the data. That being said, the example models include objective functions that are made up to illustrate the SEPOMF concept. These example objective functions were not developed using actual data and therefore a comparison of these to the available survey data has limited value.

Therefore, no questions are asked of validators.

G.4.2.12 Behavior – Surprise behavior

Question from Richardson and Pugh: Does the model under some test circumstances produce dramatically unexpected behavior, not observed in the real system?

Analysis: The SEPOMF models are optimization models. Optimization models may find solutions that are counter to the expectations of the analysts. When this is the case, analysts should carefully scrutinize the outputs of the optimization model to be certain that the recommendations are truly optimal. One valuable question analysts should consider is if it is possible that the optimization model found a local optima versus the global optima. However, for the example models that are generated using the SEPOMF, the parameters are made up explicitly to produce results that appear unsurprising. Therefore, it is impossible to assess with the available data whether models generated using the SEPOMF are capable of producing unexpected results.

Therefore, no questions are asked of validators.

G.4.2.13 Behavior – Extreme conditions

Question from Richardson and Pugh: Does the model behave reasonably under extreme conditions or extreme policies, even ones that have never been observed in the real system?

Analysis: The example models feature objective functions that are made up. Because of this, it is impossible to define what extreme conditions mean in the context of examples not based on actual data.

Therefore, no questions relating to this item are asked of validators.

G.4.2.14 Behavior – Statistical tests

Question from Richardson and Pugh: Does the model output have statistically like data for the real system?

Analysis: The example models feature objective functions that are made up. Because of this, comparing model output with the survey results is meaningless.

Therefore, no questions are asked of validators.

G.4.2.15 Structure – Appropriateness of structure

Question from Richardson and Pugh: Is the size of the model, its simplicity or complexity, and its level of aggregation or richness of detail appropriate for the audience for the study?

Analysis: A closely analogous question should be asked of the modeling framework.

Question for Validators: Is the level of simplicity or complexity and level of aggregation or richness of detail in the SEPOMF appropriate to allow a decision maker to maximize the probability of contract award given a specific budget? If not, please explain.

G.4.2.16 Behavior – Counterintuitive Behavior

Question from Richardson and Pugh: In response to some policies, does the model exhibit behavior that at first contradicts intuitions and later, with the aid of the model, is seen as a clear implication of the structure of the system?

Analysis: The example models feature objective functions that are made up. Because of this, it is of little value to determine if the example models exhibit counterintuitive behavior.

Therefore, no questions are asked of validators.

G.4.2.17 Behavior – Generation of Insights

Question from Richardson and Pugh: Is the model capable of generating new insights, or at least the feeling of new insights, about the nature of the problem addressed and the system within which it arises?

Analysis: The example models feature objective functions that are made up. Because of this, it is meaningless to attempt to gain insight about potential solutions from these example models.

Therefore, no questions are asked of validators.

G.4.2.18 General Questions about the SEPOMF

The final question is asked to provide validators an opportunity to express any opinions about the SEPOMF that they did not have an opportunity to express in the other more directed questions that were asked relating to the SEPOMF.

Question for Validators: Do you have any other comments or concerns regarding the modeling framework?

G.4.2.19 Model Tests Adapted to the SEPOMF

Table G.2 provides an overview of the relevant model test related questions customized to be applicable to the SEPOMF. In some cases, a single model test corresponds to multiple questions. As is discussed in the previous sections, many of the model tests are not applicable. The letters N/A are placed by any model tests that are not applicable.

G.5 Input from Validators About the SEPOMF

Four validators with industry experience reviewed the SEPOMF and provided feedback. The four validators have experience in four different organizations. Two validators have industry experience from the same organization, but one validator's industry experience spans two organizations. The following set of tables (Table G.3 through Table G.6) is organized by validator and provides responses to each question. The first column in these tables includes the questions that the validators were asked. The same set of questions is asked of all four validators. The second column contains the validators' responses to the questions in the first column. Responses are copied exactly from the form that the validators populated using the spelling, grammar, underlining, bolding, italicizing and punctuation originally used by the validator. This is done to fully communicate the responses. Blanks spaces in the 2nd column indicate that the validator did not respond to the particular item.

In cases where there is specific information in the responses that indicates association with a particular organization, project, system or customer, that identifying information is reserved to protect the anonymity of the participant. In such cases, the text "{Reserved}" appears where information is being withheld to protect the anonymity of the validator. After the input of each validator is presented, an analysis of this input will be conducted. For each question, responses from different validators will be presented alongside each other and discussed. Resolutions will be offered to the validator inputs.

G.6 Resolutions to Validator Comments

The inputs from validators related to the SEPOMF in Tables G.3 through G.6 are compared in Table G.7. This is done so that for each question in the validation package, the responses may be analyzed and addressed as a whole. Only questions 7-13 are analyzed in Table G.7. This subset of questions pertains to the SEPOMF as opposed to the background and qualifications of the validators. The responses from all validators to the first few questions clearly establish that each of the four validators meets the qualifications to validate, and therefore should have their input considered. The first column in Table G.7 is a repeat of each question. The second column contains the responses of the validators. For each row in the first column, there are four rows in the second column. Each row contains the response of a particular validator to the question. For each question, the first sub-row contains the response from Validator #1, the second sub-row contains the response from Validator #2, the third sub-row contains the response from Validator #3, and the fourth sub-row contains the response from Validator #4. As with Table G.3 through Table G.6, the responses in the second column are exactly as provided by the respondent, preserving spelling, grammar, underlining, bolding, italicizing and punctuation. The third column presents the resolutions to the validator comments. For each question, these resolutions are all presented in one cell. In other words, there is a one-to-one map between questions and cells in column 3. However, there is not a one-to-one mapping between validator feedback and resolutions. In some cases, a single validator will identify multiple issues in responses to a single question. In other cases, multiple validators may raise similar concerns that only need to be addressed once. The resolutions in the third column are prefaced by a short phrase to convey the content being addressed. Where there is more than one resolution to a particular question, each resolution is numbered to help the reader understand the breaks between topics.

Based upon the complete set of responses for each question, resolutions are provided. These resolutions were formulated by comparing the different validator responses to see if common issues were raised by more than one validator, by evaluating the relevance of each comment with respect to the larger research goals of the study and by determining if each issue relates to a fundamental shortcoming of the research or is simply caused by the limited subset of information about the research that was provided to validators at the time of validation. In some cases, resolutions describe modifications to this dissertation or the SEPOMF model definition. In some cases, the resolutions highlight particular features of the model that are believed to address the validators concerns but may not have been clearly communicated by the materials given validators to review. In other cases, the resolutions are simply responses to the validator comments.

Table G.2 Relevant Test Questions for the SEPOMF

Activity Type	Test	Relevant Questions
Consistency	Structure - Dimensional consistency	N/A
	Structure - Extreme condition tests in equations	N/A
	Structure - Boundary structural adequacy tests	1.) Are there important variables that are missing in the SEPOMF formulation that should be included? If so, please explain. 2.) Are there variables included in the SEPOMF formulation that are extraneous and should be removed? If so, please specify which variables and explain why you believe that they are extraneous.
	Behavior – Parameter (in)sensitivity	N/A
	Behavior – Structural (in)sensitivity	N/A
Suitability	Structure - Face validity	1.) Do the equations and inequalities (e.g., the constraints) presented in the mathematical programming formulation of the optimization model appear to reasonably represent the variables and relationships related to the use of systems engineering on proposals? If not, please describe issues that you see. 2.) Do you believe that a reasonable fit exists between the variables and structure of the SEPOMF and the essential characteristics of the use of systems engineering on proposals?
	Structure - Parameter validity	Are the variables presented in the mathematical programming formulation of the optimization model recognizable in terms of the use of systems engineering on proposals? If not, please explain.
	Behavior – Replication of reference modes	N/A
	Behavior – Surprise behavior	N/A
	Behavior – Extreme conditions	N/A

Table G.2—*Continued*

	Behavior – Statistical tests	N/A
Model Utility and Effectiveness	Structure – Appropriateness of structure	Is the level of simplicity or complexity and level of aggregation or richness of detail in the SEPOMF appropriate to allow a decision maker to maximize the probability of contract award given a specific budget? If not, please explain.
	Behavior – Counterintuitive Behavior	N/A
	Behavior – Generation of insights	N/A

Table G.3 Inputs from Validator #1

Question	Validator Response
1. Name	Reserved (Interview on 07/30/2013, Packet Received 09/03/13)
2. Have you used systems engineering on proposals for complex systems? If so, please provide a brief description.	Yes. In my 29 year career in the defense industry I have been involved in a number of proposal efforts, all as an engineer performing engineering estimates and technical content.
3. Approximately how many proposal efforts have you contributed to from the supplier perspective where complex systems are being engineered? If you have been involved in the role of supporting one or more proposal efforts from the acquirer perspective, please do not include those efforts in this estimate.	Approximately 8 proposals
4. Do you have expertise in systems engineering processes and why do believe this?	My undergraduate engineering program included some courses in systems engineering; however my main experience began at {Reserved} in 1985 when I was asked to join a group of senior engineers working proposals and system integration efforts. They effectively tutored me in application of systems engineering techniques for system design and estimation.
5. Do you understand the basic concepts of optimization modeling and why do you believe this?	As an engineer and physicist (I have two Bachelor's degrees, one in Physics, the other in Electrical Engineering) I was exposed to mathematical concepts on optimization; I have acquired more detailed knowledge as a member of Operations Analysis engineering groups at {Reserved} and {Reserved}. I have applied some optimization techniques in modeling and simulation as part of that work.
6. Do you understand the basic concepts of regression modeling in statistics and why do you believe this?	Again, in my work in Operations Analysis I have needed to learn statistical techniques including regression techniques.
7. Are the variables presented in the mathematical programming formulation of the optimization model (in slide 37) recognizable in terms of the use of systems engineering on proposals? If not, please explain.	yes

Table G.3—*Continued*

8. Do the equations and inequalities (e.g., the constraints) presented in the mathematical programming formulation of the optimization model (in slide 37) appear to reasonably represent the variables and relationships related to the use of systems engineering on proposals? If not, please describe issues that you see.	The definitions are clear; however the slide does not show u or t as they contribute to P(x). I assume that x, u, and t are contributors to total *x* in $\hat{\mathrm{P}}(x)$
9. Are there important variables that are missing in the SEPOMF formulation that should be included? If so, please explain.	It probably should at least be noted that B has a non-asymptotic relationship to P. That is, increasing proposal budget infinitely does not necessarily increase probability of contract award to unity. It is also entirely possible, even likely, that increasing number of labor hours beyond some point may become self-defeating, resulting in a lower Pwin.
10. Are there variables included in the SEPOMF formulation that are extraneous and should be removed? If so, please specify which variables and explain why you believe that they are extraneous.	No.
11. Is the level of simplicity or complexity and level of aggregation or richness of detail in SEPOMF appropriate to allow a decision maker to maximize the probability of contract award given a specific budget? If not, please explain.	A serious confounding factor exists in attempting to quantify pre-existing customer relationship. It is possible that a very small proposal team will have a very high P in areas where a good customer relationship has already been established, due to the “trust factor”. This may be accounted for by using specific historical data in training the model, so long as application of the model remains in the business area defined by the historical data.
12. Do you believe that a reasonable fit exists between the variables and structure of SEPOMF and the essential characteristics of the use of systems engineering on proposals?	Yes
13. Do you have any other comments or concerns regarding the modeling framework?	This appears to be a reasonable attempt to quantify an area of much uncertainty.
14. May I follow up with you regarding the answers that you have provided if I have any further questions or feel that I could benefit from further clarification?	Certainly!

Table G.4 Inputs from Validator #2

Question	Validator Response
1. Name	Reserved (Interview on 08/07/2013, Packet Received 08/08/13)
2. Have you used systems engineering on proposals for complex systems? If so, please provide a brief description.	Many times on proposals with wide range of size, complexity, and scope. In many kinds of industries and endeavors from retail, medical defense, industrial, manufacturing, R&D, etc.
3. Approximately how many proposal efforts have you contributed to from the supplier perspective where complex systems are being engineered? If you have been involved in the role of supporting one or more proposal efforts from the acquirer perspective, please do not include those efforts in this estimate.	50
4. Do you have expertise in systems engineering processes and why do believe this?	Performed S.E. and P.M. on many programs, consulted with over 150 companies doing programs where many were employing S. E. and have taught S. E. at the Graduate level.
5. Do you understand the basic concepts of optimization modeling and why do you believe this?	Yes Have had optimization performed on many of my programs in various forms and methods.
6. Do you understand the basic concepts of regression modeling in statistics and why do you believe this?	Yes have used it personally on many of my programs.
7. Are the variables presented in the mathematical programming formulation of the optimization model (in slide 37) recognizable in terms of the use of systems engineering on proposals? If not, please explain.	Yes
8. Do the equations and inequalities (e.g., the constraints) presented in the mathematical programming formulation of the optimization model (in slide 37) appear to reasonably represent the variables and relationships related to the use of systems engineering on	Yes

Table G.4—*Continued*

proposals? If not, please describe issues that you see.	
9. Are there important variables that are missing in the SEPOMF formulation that should be included? If so, please explain.	Make sure the ground rules and assumptions about where the study starts in the proposal timeline and what variables are already considered fixed and what are left open to optimize. Also make sure it is explained that this will only result in a framework for a method, not an actual method because there is not enough raw proposal data available to get to a real closed form method yet.
10. Are there variables included in the SEPOMF formulation that are extraneous and should be removed? If so, please specify which variables and explain why you believe that they are extraneous.	No these are ok provided the ground rules and assumptions about the starting point are clear. Also reference our discussion about including risk/opportunity assessment and mitigation in the ground rules story.
11. Is the level of simplicity or complexity and level of aggregation or richness of detail in SEPOMF appropriate to allow a decision maker to maximize the probability of contract award given a specific budget? If not, please explain.	This is a very simplified approach with very limited data going into really creating a "method'. It is after all only a framework for a method and not really a truly useful method for the user until enough real proposal data is researched to derive the inputs required to complete the framework into a method.
12. Do you believe that a reasonable fit exists between the variables and structure of SEPOMF and the essential characteristics of the use of systems engineering on proposals?	Yes
13. Do you have any other comments or concerns regarding the modeling framework?	Don't really know how what you're going to come up with is really going to be that useful until enough data is mined to fill in the framework but I guess it's at least a start to getting to a method someday.
14. May I follow up with you regarding the answers that you have provided if I have any further questions or feel that I could benefit from further clarification?	Yes

Table G.5 Inputs from Validator #3

Question	Validator Response
1. Name	Reserved (Interview on 08/16/2013, Packet Received 08/29/13)
2. Have you used systems engineering on proposals for complex systems? If so, please provide a brief description.	Yes. Case 1: In order to determine an accurate set of technical requirements for a large command & control system, a complete functional analysis was accomplished and used to provide detailed estimates of hardware and software capacity; this was followed by a architecture synthesis process to select the optimum proposed configuration. Case 2: Synthesis of a logical architecture led to the synthesis of an optimum physical architecture.
3. Approximately how many proposal efforts have you contributed to from the supplier perspective where complex systems are being engineered? If you have been involved in the role of supporting one or more proposal efforts from the acquirer perspective, please do not include those efforts in this estimate.	10
4. Do you have expertise in systems engineering processes and why do believe this?	Yes. 1. Practice of systems engineering activities over 32 year career in aerospace engineering. 2. Graduate level teaching of systems engineering processes and practices for 8 cohorts in thee {Reserved} PROGRAM
5. Do you **understand** the basic concepts of optimization modeling and why do you believe this?	Yes. Optimization modeling is a fundamental element of aerospace system sizing.
6. Do you understand the basic concepts of regression modeling in statistics and why do you believe this?	Yes. Use of regression modeling for data analysus.
7. Are the variables presented in the mathematical programming formulation of the optimization model (in slide 37) recognizable in terms of the use of systems engineering on proposals? If not, please explain.	Yes.
8. Do the equations and inequalities (e.g., the	Yes.

Table G.5—*Continued*

constraints) presented in the mathematical programming formulation of the optimization model (in slide 37) appear to reasonably represent the variables and relationships related to the use of systems engineering on proposals? If not, please describe issues that you see.	
9. Are there important variables that are missing in the SEPOMF formulation that should be included? If so, please explain.	None that are immediately obvious.
10. Are there variables included in the SEPOMF formulation that are extraneous and should be removed? If so, please specify which variables and explain why you believe that they are extraneous.	No.
11. Is the level of simplicity or complexity and level of aggregation or richness of detail in SEPOMF appropriate to allow a decision maker to maximize the probability of contract award given a specific budget? If not, please explain.	Yes. However, adequate explanations in layman's terms should be included with the model since not all future users will have the adequate background to understand them in context.
12. Do you believe that a reasonable fit exists between the variables and structure of SEPOMF and the essential characteristics of the use of systems engineering on proposals?	Yes
13. Do you have any other comments or concerns regarding the modeling framework?	The framework is an excellent foundation to future application.
14. May I follow up with you regarding the answers that you have provided if I have any further questions or feel that I could benefit from further clarification?	Yes.

Table G.6 Inputs from Validator #4

Question	Validator Response
1. Name	Reserved (Interview on 08/17/2013, Packet Received 09/01/13)
2. Have you used systems engineering on proposals for complex systems? If so, please provide a brief description.	I have applied systems engineering to propose development of several {Reserved} flight missions that were valued between $10M and $100M.
3. Approximately how many proposal efforts have you contributed to from the supplier perspective where complex systems are being engineered? If you have been involved in the role of supporting one or more proposal efforts from the acquirer perspective, please do not include those efforts in this estimate.	10
4. Do you have expertise in systems engineering processes and why do believe this?	I have over twenty years of {Reserved} service, hold a systems engineering certificate from the {Reserved}, and successfully completed a number of systems engineering assignments for various projects authorized by the {Reserved} Program, {Reserved} Program, {Reserved} Program, and numerous {Reserved} Programs.
5. Do you understand the basic concepts of optimization modeling and why do you believe this?	I have a situational understanding of systems optimization as a control that collects feedback, and measures variation, and makes changes to produce a desired outcome.
6. Do you understand the basic concepts of regression modeling in statistics and why do you believe this?	My understanding of regression analysis is basic. My understanding is that it is a statistical technique for studying relationships between dependent and independent variables, useful to predict the value of the dependent variable, or estimate the effect of some explanatory variable on the dependent variable.
7. Are the variables presented in the mathematical programming formulation of the optimization model (in slide 37) recognizable in terms of the use of systems engineering on proposals? If not, please explain.	My experience with this technique is limited. The variables used in terms of systems engineering are recognizable but limited. It is my opinion that the program over simplifies proposal optimization.
8. Do the equations and inequalities (e.g., the constraints) presented in the mathematical programming formulation of the optimization	

Table G.6—*Continued*

model (in slide 37) appear to reasonably represent the variables and relationships related to the use of systems engineering on proposals? If not, please describe issues that you see.	
9. Are there important variables that are missing in the SEPOMF formulation that should be included? If so, please explain.	Important variables appear to be missing in the SEPOMF formulation. Proposal optimization should address the following; compliance with customer requirements, quality of key differentiators of proposition, strength of customer relationship, awareness of competitors' strength, relative competitive position of the proposing entity, awareness of political, economic, & environmental conditions, significance of bid & proposal investment, level of concept maturity, schedule & cost, availability & quality needed procurements, workforce, equipment, & facilities, financial & technical leverage, partners to the endeavor, success criteria, risk management, quality standards, resource needs, awareness of stakeholders expectations, business case quality, prominence of key personnel, past performance of the proposing organization….
10. Are there variables included in the SEPOMF formulation that are extraneous and should be removed? If so, please specify which variables and explain why you believe that they are extraneous.	
11. Is the level of simplicity or complexity and level of aggregation or richness of detail in SEPOMF appropriate to allow a decision maker to maximize the probability of contract award given a specific budget? If not, please explain.	
12. Do you believe that a reasonable fit exists between the variables and structure of SEPOMF and the essential characteristics of the use of systems engineering on proposals?	
13. Do you have any other comments or concerns regarding the modeling framework?	Questions 10, 11, and 12 are unanswered because of my opinion that the model oversimplifies what is sought after, an "optimized proposal", or to generally assess proposals competitiveness. I do wish you success in your

Table G.6—*Continued*

	pursuit.
14. May I follow up with you regarding the answers that you have provided if I have any further questions or feel that I could benefit from further clarification?	Yes.

Table G.7 Validation Comment Resolution Matrix

Question	Validator Response	Resolution
7. Are the variables presented in the mathematical programming formulation of the optimization model (in slide 37) recognizable in terms of the use of systems engineering on proposals? If not, please explain.	yes	No action necessary – All validators agree that the variables are recognizable in terms of the use of systems engineering on proposals.
	Yes	
	Yes.	
	My experience with this technique is limited. The variables used in terms of systems engineering are recognizable but limited. It is my opinion that the program over simplifies proposal optimization.	
8. Do the equations and inequalities (e.g., the constraints) presented in the mathematical programming formulation of the optimization model (in slide 37) appear to reasonably represent the variables and relationships related to the use of systems engineering on proposals? If not, please describe issues that you see.	The definitions are clear; however the slide does not show u or t as they contribute to P(x). I assume that x, u, and t are contributors to total *x* in $\hat{\mathrm{P}}(x)$	Role of constraints - The following has been added to the discussion in Section 3.5.3 to help clarify the role of u and t in determining an optimal solution: "Observe that *u* and *t* are not factors in *P*, but rather proposal-effort specific constraints that help define the feasible region. Because the values for *u* and *t* constrain the options, they in fact can impact the optimal solution even if not represented directly in the objective function."
	Yes	
	Yes.	
9. Are there important variables that are missing in the SEPOMF formulation that should be included? If so, please explain.	It probably should at least be noted that B has a non-asymptotic relationship to P. That is, increasing proposal budget infinitely does not necessarily increase probability of contract award to unity. It is also entirely possible, even likely, that increasing number of labor hours beyond some point may become self-defeating, resulting in a lower Pwin.	1. Relationship between *B* and *P* – The following text has been added to Section 3.5.3: "Logistic regression functions never assume a value of 0 or 1, but approach these values asymptotically. This property of logistic regression functions is well-aligned with the reality of making investments in systems engineering labor on proposals. Even if the available budget for systems engineering labor approaches infinity,
	Make sure the ground rules and assumptions about where the study starts in the proposal timeline and what variables are already	

Table G.7—*Continued*

<table>
<tr><td rowspan="3"></td><td>considered fixed and what are left open to optimize. Also make sure it is explained that this will only result in a framework for a method, not an actual method because there is not enough raw proposal data available to get to a real closed form method yet.</td><td rowspan="3">there is generally no guarantee of a contract award.”
2. To clarify when the study starts in the proposal timeline of interest - Figure 3.1 is provided.
3. To clarify assumptions and ground rules – Additional preconditions in the use case definition in Table 3.5 have been added to more fully explain ground rules for when it is appropriate to use the SEPOMF.
4. To clarify the fact data was not included in the model - Discussions throughout the dissertation (e.g., Chapter 1, Chapter 3, Chapter 4, Appendix F) discuss that the SEPOMF is a framework and that the data collected by the survey was not sufficient to actually develop objective functions.
5. The role of fixed variables in the model –The SEPOMF is sufficiently flexible so that whatever subset of variables is relevant can be included in the model. The set of variables or factors explored in the survey in this dissertation is a subset of possible variables or factors that can be used in the SEPOMF. The validator feedback includes some very promising options for variables to potentially be explored in future research. One factor not explored in detail in the survey that is suggested by multiple validators is level of risk. Exploring level of risk as a factor is now</td></tr>
<tr><td>None that are immediately obvious.</td></tr>
<tr><td>Important variables appear to be missing in the SEPOMF formulation. Proposal optimization should address the following; compliance with customer requirements, quality of key differentiators of proposition, strength of customer relationship, awareness of competitors’ strength, relative competitive position of the proposing entity, awareness of political, economic, & environmental conditions, significance of bid & proposal investment, level of concept maturity, schedule & cost, availability & quality needed procurements, workforce, equipment, & facilities, financial & technical leverage, partners to the endeavor, success criteria, risk management, quality standards, resource needs, awareness of stakeholders expectations, business case quality, prominence of key personnel, past performance of the proposing organization….</td></tr>
</table>

Table G.7—*Continued*

		explicitly discussed as potential future work in Section 5.4.
10. Are there variables included in the SEPOMF formulation that are extraneous and should be removed? If so, please specify which variables and explain why you believe that they are extraneous	No.	Risk/opportunity management in the ground rules – The recommendation is made to address risk as a factor warranting future study. Risk is not explicitly included in the preconditions or ground rules for using a SEPOMF model. It is assumed that organizations may have considered risks and opportunities when deciding to pursue the contract. However, some organizations may not use level or risk or opportunity as a go/no-go criteria for pursuing a particular contract.
	No these are ok provided the ground rules and assumptions about the starting point are clear. Also reference our discussion about including risk/opportunity assessment and mitigation in the ground rules story	
	No.	
11. Is the level of simplicity or complexity and level of aggregation or richness of detail in SEPOMF appropriate to allow a decision maker to maximize the probability of contract award given a specific budget? If not, please explain.	A serious confounding factor exists in attempting to quantify pre-existing customer relationship. It is possible that a very small proposal team will have a very high P in areas where a good customer relationship has already been established, due to the “trust factor”. This may be accounted for by using specific historical data in training the model, so long as application of the model remains in the business area defined by the historical data.	1. Pre-existing customer relationship – The survey data analysis has identified level of customer satisfaction on previous or ongoing contract efforts as highly correlated with an increased probability of contract awards. The SEPOMF allows for any number of additional variables beyond the decision variables explored already (including ones pertaining to the pre-existing customer relationship) to be represented in the model. Treating level of trust between proposal team and customer as a factor in the model is an alternative to calibrating the model only with data from a similar customer base, yet would allow the level of trust to be considered when estimating probability of contract award. 2. Laymen’s terms –This dissertation is
	This is a very simplified approach with very limited data going into really creating a “method’. It is after all only a framework for a method and not really a truly useful method for the user until enough real proposal data is researched to derive the inputs required to complete the framework into a method.	
	Yes. However, adequate explanations in layman’s terms should be included with the	

Table G.7—*Continued*

	model since not all future users will have the adequate background to understand them in context.	written for audiences with significant depth of understanding of analytical models. It is recommended that if an organization attempts to leverage this research to develop such a DSS, they include at least one person with technical depth in the areas of regression analysis and optimization. Section 1.9 has been added to clarify who the intended audience is for the dissertation and what skill sets will be required to leverage the information in this dissertation.
12. Do you believe that a reasonable fit exists between the variables and structure of SEPOMF and the essential characteristics of the use of systems engineering on proposals?	Yes.	No action required – No validators identified discrepancies between the variables and structure of the SEPOMF and the characteristics of the use of systems engineering on proposals.
	Yes	
	Yes	
13. Do you have any other comments or concerns regarding the modeling framework?	This appears to be a reasonable attempt to quantify an area of much uncertainty.	1. Developing real models with data – Developing a model based on real data will provide an opportunity to prove this concept as well as serve as a more concrete example that can be more easily replicated. Gathering an adequate quantity of data to use to develop a DSS is recommended as future work in Chapter 5. 2. Oversimplification - This framework is very simplistic by design. First, this framework is designed to maximize the probability of contract award when in reality this is just one of a number of important objectives that can be
	Don't really know how what you're going to come up with is really going to be that useful until enough data is mined to fill in the framework but I guess it's at least a start to getting to a method someday.	
	The framework is an excellent foundation to future application.	
	Questions 10, 11, and 12 are unanswered because of my opinion that the model oversimplifies what is sought after, an "optimized proposal", or to generally assess proposals competitiveness. I do wish you success in your pursuit.	

Table G.7—*Continued*

		considered [Smartt and Ferreira, 2012c]. Also, this framework is defined to exclusively focus on how to invest resources for systems engineering labor on a proposal. While the model provides a way to account for other important variables (e.g., level of customer satisfaction on previous or ongoing contract efforts), the model as defined in this dissertation does not explicitly describe a way for any of these other important variables to be treated as decision variables. Noting these limitations, somebody with an advanced level of understanding of applied mathematics and statistics concepts should be able to extend this model in a number of different ways with a limited effort.

References

T. K. Abdel-Hamid and S. E. Madnick, Software project dynamics: An integrated approach, Prentice-Hall, Englewood Cliffs, New Jersey, 1991.

T.K. Abdel-Hamid, A multi-project perspective of single project dynamics, Journal of Systems and Software 22(1) (1993), 151-165.

V. Agrawal and M. Ferguson, Bid-response models for customized pricing, Journal of Revenue & Pricing Management 6(3) (2007), 212-228.

A. Agresti, An introduction to categorical data analysis, John Wiley & Sons, New York, NY, 1996.

P.L. Alreck and R.B. Settle, The survey research handbook, McGraw-Hill, Chicago, IL, 1995.

ANSI, Processes for engineering a system EIA-632, 1999.

D. Arnott and G. Dodson, “Decision support system failures,” in F. Burnstein and C.W. Holsapple (Editors), Handbook on Decision Support Systems Volume 1, Springer-Verlag, Berlin, 2008.

K. J. Arrow, Existence of equilibrium for a competitive economy, Econometrica 22(3) (1954), 265-290.

K.J. Arrow, An extension of basic theorems of classical welfare economics, Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, CA, 1951.

S. Bayer and D. Gann, Balancing work: Bidding strategies and workload dynamics in project-based professional service organization, System Dynamics Review 22(3) (2006), 185-211.

B.A. Beemer and D.A. Gregg, “Advisory systems to support decision making” in F. Burnstein and C.W. Holsapple (Editors), Handbook on Decision Support Systems Volume 1, Springer-Verlag, Berlin, 2008.

M. Bertolini, M. Braglia, and G. Carmigna, Application of the AHP methodology in making a proposal for a public work contract, International Journal of Project Management 24(5) (2006), 422-430.

BLS, Bureau of Labor Statistics, www.bls.gov, last accessed September 10, 2013.

B. Boehm, Software engineering economics, Prentice Hall, Englewood Cliffs, NJ, 1981.

B. Boehm and J.A. Lane, Using the incremental commitment model to integrate system acquisition, systems engineering, and software engineering, CrossTalk 20(10) (2007), 4-9.

B. Boehm and S. Sheard, What is return on investment for systems engineering?, INCOSE Insight 13(3) (2010), 44-47.

B. Boehm, R. Valerdi, and E. Honour, The ROI of systems engineering: Some quantitative results for software-intensive systems, The Journal of Systems Engineering 11(3) (2008), 221-234.

P.D. Boughton, The competitive bidding process: Beyond probability models, Industrial Marketing Management 16(2) (1987), 87-91.

F. Burnstein and C.W. Holsapple, Handbook on decision support systems volume 1, Springer-Verlag, Berlin, 2008.

E. Cagno, F. Caron and A. Perego, Multi-criteria assessment of the probability of winning in the competitive bidding process, International Journal of Project Management 19(6) (2001), 313-324.

CapturePlanning, Captureplanning.com, http://www.captureplanning.com, last accessed September 9, 2013.

W.E. Carnes, Knowledge management: a state of the practice, IEEE 7th Human Factors

Meeting, Scottsdale, AZ, 2002.
C. Carr, K. Kolehmainen and F. Mitchell, Strategic investment decision making practices: A contextual approach, Management Accounting Research 21(3) (2010), 167-184.
M. Chan, C. Wong, B. Cheung and G. Tang, Genetic algorithms in multi-stage asset allocation system, IEEE International Conference on Systems, Man and Cybernetics, Yasmine Hammarrett, Tunisia, 2002, pp. 1-6.
Y. K. Che and I. Gale, Optimal design of research contests, The American Economic Review 93(3) (2003), 646-671.
W.G. Cochran, Some methods of strengthening the common Chi-square tests, Biometrics 10(4) (1954), 417-451.
A. Cockburn, Writing effective use cases, Addison Wesley, Boston, MA, 2001.
J. Collier, Effective long range business planning, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1968.
P.D. Cook, W.M. Duckworth, M.S. Kaiser, K. Koehler, W.Q. Meeker and W.R. Stephenson, Binary response and logistic regression analysis, http://www.public.iastate.edu/~stat415/stephenson/stat415_chapter3.pdf, 2001, last accessed September 10, 2013.
Cytel, Logxact, http://www.cytel.com/software/logxact, last accessed September 10, 2013.
G.B. Dantzig and G. Infanger, Multi-stage stochastic linear programs for portfolio optimization, Annals of Operations Research, 45(1) (1993), 59-76.
DAU, Defense acquisition guidebook, Defense Acquisition University, http://www.dua.mil/dag, 2012, last accessed April 5, 2013.
DAU, Defense acquisition university, Defense Acquisition University, http://www.dua.mil/default.aspx, 2013, last accessed April 5, 2013.
H. L. Davidz and D. J. Nightingale, Enabling systems thinking to accelerate the development of senior systems engineers, The Journal of Systems Engineering 11(1) (2008), 1-14.
R. de Neufville, Real options: dealing with uncertainty in systems planning and design, Integrated Assessment 4(1) (2003), 26-34.
A. Dean and D. Voss, Design and analysis of experiments, Springer-Verlag, New York, NY, 1999.
C. Demir, S.H. Gultekin and B. Yener, Learning the topological properties of brain tumors, IEEE/ACM Transactions on Computational Biology and Bioinformatics 2(3) (2005), 262-270.
D.A. Dillman, “Mail and other self-administered questionnaires,” in P.H. Rossi, J.D. Wright and A.B. Anderson (Editors), Handbook of Survey Research, Academic Press Inc., New York, NY, 1983.
DoD, Manual for the operation of the Joint Capability Integration and Development System (JCIDS), United States Department of Defense, 2011.
DoD, Systems engineering guide for systems of systems, United States Department of Defense, 2008.
C. Doelling and S. Ferreira, Advancing the integration of systems engineering and knowledge management, Systems Research Forum 4(1) (2010), 33-44.
A. Efron and R.J. Tibshirami, An introduction to the bootstrap, Chapman and Hall, New York, NY, 1993.
H. Eisner, Essentials of project and systems engineering management 2nd edition, John Wiley and Sons, New York, NY, 2002.
J. Elm, The business case for systems engineering study: Assessing project performance

with sparse data, Software Engineering Institute (SEI) at Carnegie Mellon University Special Report CMU/SEI-2012-SR-010, 2012.

J. Elm, D.R. Goldenson, K.E. Enam, N. Donatelli and A. Neisa, A survey of systems engineering effectiveness, Software Engineering Institute (SEI) at Carnegie Mellon University, CMU/SEI-2007-SR-014, 2007.

H. Eriksson and M. Penker, UML toolkit, John Wiley and Sons, New York, NY, 1998.

FAA, National airspace systems engineering manual, Federal Aviation Administration, 2006.

S. Ferreira, Measuring the effects of requirements volatility on software development Projects, Ph.D. Dissertation, Arizona State University, Tempe, AZ, USA, 2002.

A. Fink, The survey handbook, Sage Publications, Thousand Oaks, CA, 1995.

D. Firth, Bias reduction of maximum likelihood estimates, Biometrika 80(1) (1993), 27-38.

E.H. Forman and S.I. Gass, The analytical hierarchy process: An exposition, Operations Research 49(4) (2001), 469-486.

J. Fortune, R. Valerdi, B. Boehm and F.S. Settles, Estimating systems engineering reuse, Conference on Systems Engineering Research 2009 (CSER 2009), Loughborough, UK, 2009, pp. 1 – 10.

F.J. Fowler, Survey research methods: second edition, SAGE Publications, Newbury Park, CA, 1993.

M. Frankel, “Sampling theory”, in P.H. Rossi, J.D. Wright and A.B. Anderson (Editors), Handbook of survey research, 1983, 21-69.

D. Gann and A. Salter, Learning and innovation management in project-based, service-enhanced firms, International Journal of Innovation Management 2(4) (1998), 431-454.

G. A. Garrett and R. J. Kipke, The capture management lifecycle: Winning more business, CCH Incorporated, Chicago, 2003.

D.H. Gensch, N. Aversa and S.P. Moore, A choice-modeling market information system that enable ABB Electric to expand its market, Interfaces 20(1) (1990), 6-25.

G.A. Gorry and M.S. Scott Morton, A framework for information management systems, Sloan Management Review 13(1) (1971), 55-70.

D.R. Greer, L.J. Black and R.J. Adams, Improving inter-organizational baseline alignment in large space system development programs, 2006 IEEE Aerospace Conference, Big Sky, MT, 2006, pp 1 -14.

A. Grobler, T. Lopsinger, M. Stotz and H.. Wormer, Analyzing price and product strategies with a comprehensive system dynamics model – a case study from the capital goods industry, Journal of Business Research 61(11) (2008), 1136-1142.

F.E. Harrell, K.L. Lee, D.B. Matchar and T.A. Reichert, Regression models for prognostic prediction: advantages, problems, and suggested solutions, Cancer Treatment Reports 69(10) (1985), 1071 – 1077.

W.W. Hines and D.C. Montgomery, Probability and statistics in engineering and management science, The Ronald Press Company, New York, 1972.

O. Hosny and A. Elhakeem, Simulating the winning bid: a generalized approach for optimum markup estimation, Automation in Construction 22(1) (2012), 357-367.

D.C. Howell, Treatment of missing data, http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Missing.html, 2013, last accessed April 5, 2013.

C.W. Holsapple, “DSS architectures and types” in F. Burnstein and C.W. Holsapple (Editors), Decision Support Systems Volume 1, Springer-Verlag, Berlin, 2008.

E. Honour, Sizing systems engineering activities to optimize return on investment, INCOSE International Symposium 2011, Denver, CO, 2011, pp. 1-13.

T. Hutzschenreuter and S. Israel, A review of empirical research on dynamic competitive strategy, International Journal of Management Reviews 11(4) (2009), 421-461.
G. Iarossi, The power of survey design, The World Bank, Washington, DC, 2006.
IBM, IBM SAS Bootstrapping, http://www-01.ibm.com/software/analytics/spss/products/statistics/, 2011, last accessed June 19, 2013.
IBM, SPSS statistics, ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Bootstrapping.pdf, last accessed January 21, 2013.
Idealware, A few good online survey tools, http://www.idealware.org/articles/fgt_online_surveys.php, last accessed September 11, 2013.
IDRE, SAS annotated output: proc logistic, Institute for Digital Research and Education, University of California at Los Angeles, http://www.ats.ucla.edu/stat/sas/output/sas_logit_output.htm, 2013, last accessed April 5, 2013.
INCOSE, INCOSE website, http://www.incose.org/, last accessed September 11, 2013.
INCOSE, Systems engineering handbook version 2a: A "what to" guide for all SE practitioners, 2004.
INCOSE, Systems engineering handbook: A guide for system life cycle processes and activities version 3.2, Editor: Cecilia Haskins, 2010.
ISO/IEC, ISO/IEC-15288: systems engineering: systems lifecycle processes, International Standards Organization/IEEE Computer Society, 2008.
J. H. Kagel and D. Levin, The winner's curse and public information in common value auctions, American Economic Review 76(5) (1986), 894-920.
M.D. Kaplowitz, A comparison of web and mail survey response rates, The Public Opinion Quarterly 68(1) (2004), 94-101.
P.G.W. Keen and J.H. Scott Morton, Decision support systems: An organizational perspective, Addison-Wesley, Reading, MA, 1978.
C.T. Kelley, Iterative methods for optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.
H. P. Kerzner, Project management: A systems approach to planning, scheduling, and controlling, Hoboken, NJ, 2003.
G. Kuhn, 7 common pitfalls when managing a survey, http://rmsbunkerblog.wordpress.com/2011/01/03/7-common-pitfalls-when-managing-a-survey-marketing-in-syracuse-ny-market-research-upstate-ny-central-ny-new-york/, last accessed November 29, 2012.
M.H. Kutner, C.J. Nachtsheim, J. Neter and W. Li, Applied linear statistical models, McGraw-Hill Irwin, Boston, MA, 2005.
J. Lane, Cost model extensions to support systems engineering cost estimation for complex systems and systems of systems, 2009 Conference on Systems Engineering Research (CSER 2009), Loughborough, UK, 2009, pp. 1-8.
L.P. Leach, Critical chain project management, Artech House, Boston, MA, 2000.
H. Lewis, Bids, tenders, and proposals: Winning business through best practice, Kogan Page, London, 2003.
R.J.A. Little, Missing data adjustments in large surveys, Journal of Business & Economic Statistics 6(3) (1988), 287-296.
M.S. Litwin, How to measure survey reliability and validity, Sage Publications, Thousand Oaks, CA, 1995.
D.J. Lowe and J. Parvar, A logistic regression approach to modeling the contractor's

decision to bid, Construction Management and Economics 22(6) (2004), 643-653.
T. M. Ludden, S.L. Beal and L.B. Scheiner, Comparison of Akaike Information Criterion, the Schwarz Criterion and the f test as guides to model selection, Journal of Pharmacokinetics and Biopharmaceutics 22(5) (1994), 431 – 445.
T. Maiti and V. Pradham, A comparative study of bias corrected estimates in logistic regression, Statistical Methods in Medical Research 17(6) (2007), 621-634.
H. Markowitz, Portfolio selection, The Journal of Finance 7(1) (1952), 77-91.
M. Marzouk and O. Moselhi, A decision support tool for construction bidding, Construction Innovations 3(2) (2003), 111-124.
Mathworks, Matlab, http://www.mathworks.com/products/matlab/, last accessed June 19, 2013.
R. P. McAfee and J. McMillan, Auctions and bidding, Journal of Economic Literature 25(2) (1987), 699-738.
W.Q. Meeker and L.A. Escobar, Teaching about approximate confidence regions based on maximum likelihood estimation, The American Statistician 49(1) (1995), 48-53.
C.R. Mehta and N.R. Patel, IBM SPSS exact tests, IBM Corporation, Cambridge, MA, 2011.
Merriam-Webster, Merriam-Webster online, http://www.merriam-webster.com/, last accessed September 11, 2013.
C.L. Miller, A perception-based assessment of systems engineering effort estimation, Ph.D. Dissertation, George Washington University, Washington DC, 2012.
G.A. Milliken and D.E. Johnson, Analysis of messy data: volume 1: designed experiments, Chapman and Hall, London, 1997.
D.C. Montgomery, E.A. Peck and G.C. Vining, Introduction to linear regression analysis 5th edition, John Wiley and Sons, Hoboken, NJ, 2012.
S. Myers, Finance theory and financial strategy, Interfaces 14(1) (1984), 126-137.
NASA, National Aeronautics and Space Administration, Systems engineering handbook, 2007.
NDIA, National Defense Industrial Association website, http://www.ndia.org/Pages/Default.aspx, last accessed September 11, 2013.
G.L. Nemhauser and A.H.G. Rinooy Kan, Handbook in operations research and management science, Elsevier, Amsterdam, 1998.
D. Nickson, The bid manager's handbook, Gower Publishing Limited, Aldershot, England, 2003.
D.L. Olson, “Multi-criteria decision support” in F. Burnstein and C.W. Holsapple (Editors), Handbook on decision support systems volume 1, Springer-Verlag, Berlin, 2008.
OMG, Object Modeling Group (OMG) Unified Modeling Language (UML) version 2.4.1, 2011.
B. Oo, S.D. Drew and H.P. Lo, Applying a random coefficients logistic model to contractor’s decision to bid, Construction Management and Economics 25(4) (2007), 387-398.
OptTek, Defining decision variables, http://www.opttek.com/documentation/v62engine/OptQuest%20Engine%20Documentation/WebHelp/Defining_decision_variables.htm, last accessed June 15, 2013.
Oxford Dictionary online, http://oxforddictionaries.com, last accessed September 10, 2013.
S. Paranka, Competitive bidding strategy: A procedure for pre-bid analysis, Business

Horizons 14(3) (1971), 39-43.
A. Paul and G. Gutierrez, Simple probability models for project contracting, European Journal of Operations Research 165(2) (2005), 329-338.
P. Peduzzi, J. Concato, E. Kemper, T.R. Holford and A.R. Feinstein, A simulation study of the number of events per variable in logistic regression analysis, Journal of Clinical Epidemology 49(12) (1996), 1373 -1379.
W. Perks, On some experiments in the graduation of mortality statistics, Journal of the Institute of Actuaries 63(1) (1932), 12-57.
S.P. Philbin, Bid management: a systems engineering approach, The Journal of High Technology Management Research 19(2) (2008), 114-127.
D.J. Power, Decision support systems: Concepts and resources for managers, Quorum Books, Westport, CT, 2002.
G.P. Richardson and A. Pugh III, Introduction to system dynamics modeling, MIT Press, Cambridge, MA, 1981.
B. Roy, Problems and methods with multiple objective functions, Mathematical Programming 1(1) (1971), 239-266.
A. Rubin and E.R. Babbie, Research methods for social work, Wadsworth, Beverly, MA, 2007.
D.B. Rubin, Inferences and missing data, Biometrika 63(3) (1976), 581-592.
D.B. Rubbin and E.R. Babbie, Essential research methods for social work, Brools/Cole, Belmont, CA, 2007.
SAS, www.sas.com/?gclid=CMKv5--R8bcCFVIV7AodbB0AuA, last accessed June 19, 2013.
J.L. Schaffer, The multiple imputation FAQ page, http://sites.stat.psu.edu/~jls/mifaq.html, 2005, last accessed September 11, 2013.
SEI, Capability maturity model integration (CMMI) for development v1.3, Software Engineering Institute, http://www.sei.cmu.edu/reports/10tr033.pdf, 2010, last accessed April 5, 2013.
S. Sethi and M.E. Seligman, Optimism and fundamentalism, Psychological Science 4(4) (1993), 256-259.
S.A. Sheard and C.L. Miller, The Shangri-La of ROI, Tenth International Symposium of the International Council on Systems Engineering, Brighton, England, 2000, pp. 1 – 8.
P.B. Sheatsley, “Questionnaire construction and item writing,” in P.H. Rossi, J.D. Wright and A.B. Anderson (Editors), Handbook of Survey Research Volume 1, Academic Press, New York, NY, 1983, pp. 195-230.
J. N. Sheth, Organizational buying behavior: past performance and future expectations, Journal of Business & Industrial Marketing 11(3-4) (1996), 7-24.
R. Siegel, All-pay contests, Econometrica 77(1) (2009), 71-92.
E.S. Shtatland, E. Cain and M.B. Barton, The perils of stepwise logistic regression and how to escape them using information criterion and the output delivery system, Twenty-Sixth Annual SAS Users Group International Conference, Cary, NC, 2001, pp. 222-226.
C. Smartt and S. Ferreira, Advancing systems engineering in support of the bid and proposal process, Journal of Systems Engineering 14(3) (2011), 255-266.
C. Smartt and S. Ferreira, An analysis of systems engineering strategy, IEEE 2012 International Systems Conference, Vancouver, BC, 2012a, pp. 1-5.
C.Smartt and S. Ferreira, Applying systems engineering to survey research, 2013 Conference on Systems Engineering Research (CSER 2013), Atlanta, GA, 2013,

pp. 1102-1111.
C. Smartt and S. Ferreira, Constructing a general framework for systems engineering strategy, Journal of Systems Engineering 15(2) (2012b), 140-152.
C. Smartt and S. Ferreira, A systems engineering perspective on proposal management, IEEE Systems Journal 6(4) (2012c), 667-674.
C. Strauss and C. Stummer, Multiobjective decision support in IT-risk management, International Journal of Information Technology and Decision Making 1(2) (2002), 251-168.
SuperSurvey, Mail surveys vs. web surveys: A comparison, http://knowledge-base.supersurvey.com/mail-vs-web-surveys.htm, last accessed April 8, 2012.
A.H. Taha, Operations research: An introduction 8th edition, Pearson Prentice Hall, Upper Saddle River NJ, 2007.
W.M.K. Trochim, The research methods knowledge base, 2nd edition, http://www.socialresearchmethods.net/kb/, last accessed November 28, 2012.
E. Turban, J.E. Aronson, T. Liang and R. Sharda, Decision support and business intelligence systems 8th edition, Pearson Prentice Hall, Upper Saddle River, NJ, 2007.
N. Tweedley, Winning the bid: A manager's guide to competitive bidding, Pitman Publishing, London, 1995.
R. Valerdi, The constructive systems engineering cost model (COSYSMO), Ph.D. Dissertation, University of Southern California, Log Angeles, CA, 2005.
US, United States Government Federal Acquisition Regulations (FAR), www.acquisition.gov/far, last accessed November 4, 2013.
R. Valerdi and H. L. Davidz, Empirical research in systems engineering: Challenges and opportunities of a new frontier, The Journal of Systems Engineering 12(2) (2008), 169-181.
R. Valerdi, A.M. Ross and D.H. Rhodes, A framework for evolving systems of systems, CrossTalk 20(10) (2007b), 28-30.
R. Valerdi, J.E. Rieff, G.J. Roedler, M.J. Wheaton, and G. Wang, Lessons learned from the industrial validation of COSYSMO, International Council on Systems Engineering (INCOSE) 2007 Symposia, San Diego, CA, 2007a, pp. 1-12.
J. von Neumann and O. Morgenstern, Theory of games and economic behavior, Princeton University Press, Princeton, NJ, 1944.
G. Wang, A. Shernoff, J. Turnbridge and R. Valerdi, Toward a holistic total engineering cost model, International Council on Systems Engineering (INCOSE) 2009 Symposium, Singapore, 2009a, pp. 1-19.
G. Wang, R. Valerdi, G.J. Roedler, A. Ankrum, and J.E. Gaffney Jr., Harmonizing systems and software estimation, International Council on Systems Engineering (INCOSE) 2009 Symposium, Singapore, 2009b, pp. 1-21.
W. Wang, S. Wang, Y. Tsui and C. Hsu, A factor-based probabilistic cost model to support bid-price estimation, Expert Systems and Applications 39(5) (2012), 5358-5366.
M. Whitley, To bid or not to bid…, Engineering Management 16(5) (2006) 40-46.
J.R. Wirthlin, Identifying enterprise leverage points in defense acquisition program performance, Ph.D. Dissertation, Massachusetts Institute of Technology (MIT), Cambridge, MA, 2009.
Y. Yang, M. He, M. Li, Q. Wang and B. Boehm, Phase distribution of software development effort, 2nd International Symposium on Empirical Software Engineering and Measurement, Kairerslautern, Germany, 2008, pp. 61-69.
P. Yiftachel, D. Peled, I. Hadar and D. Goldwasser, Resource allocation among

development phases: An economic approach, International Workshop on Economics Driven Software Engineering Research, Shanghai, China, 2006, pp. 43-48.
L. Young, Interaction between sales, delivery and business functions during the proposal stage, Project Management Institute Annual Seminars & Symposium, Houston, TX, 2000, pp. 1-6.

Biographical Information

Clement Smartt holds a B.S. in Applied Mathematics from Texas A&M University, an M.S. of Applied Mathematics from Southern Methodist University, and a Ph.D. in Industrial Engineering from the University of Texas at Arlington with a research focus in systems engineering. Clement is an International Council on Systems Engineering (INCOSE) Certified Systems Engineering Professional (CSEP) since 2010. Clement is the 2012 recipient of the INCOSE Foundation award for doctoral research.

Clement is a Senior Research Scientist at the Georgia Tech Research Institute (GTRI). Previously, Clement worked as a Systems Engineer at L-3. Clement's professional career has focused on modeling and simulation, systems architecture and model based systems engineering. Clement's research interests include systems engineering, proposal management, architecture and modeling and simulation.