NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

HIGH-INFLUENCE FACTORS FOR THE TIMELINESS OF PROJECT AWARD FOR NAVY MILITARY CONSTRUCTION

by

Robert J. Thompson

September 2020

Thesis Advisor: Lyn R. Whitaker
Second Reader: Samuel E. Buttrey

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2020	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE
HIGH-INFLUENCE FACTORS FOR THE TIMELINESS OF PROJECT AWARD FOR NAVY MILITARY CONSTRUCTION

5. FUNDING NUMBERS

6. AUTHOR(S) Robert J. Thompson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

The current process that Naval Facilities Engineering Command (NAVFAC) uses to bring Military Construction (MILCON) projects from concept to contract award fails to provide reliable and timely results. This poor performance leads to the untimely delivery of critical facilities, which directly impacts warfighting and power projection capabilities. It erodes the professional reputation of NAVFAC, leaving supported commanders to question whether an essential product or service will arrive on time.

Today, there are many internal NAVFAC teams exploring process-improvement opportunities across the entire construction timeline, of which pre-award is only one piece. However, many of these efforts are limited in terms of number of projects or project factors considered.

The focus of this thesis is to analyze projects from initial documentation up to contract award. To accomplish this, we linked two widely used but independent databases to capture the complete documented life cycle of hundreds of Navy-executed MILCONs. Once linked, dozens of project factors were collected, analyzed, and used in the development of various machine-learning models to assess their influence on project award performance.

Our analysis highlights several factors among existing and newly constructed project metrics that appear to greatly influence the timeliness of project award. This collection of potentially high-influence factors can then help NAVFAC further focus its ongoing process improvements.

14. SUBJECT TERMS	**15. NUMBER OF PAGES**
MILCON, NAVFAC, pre-design, DD 1391, project documentation	101
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

HIGH-INFLUENCE FACTORS FOR THE TIMELINESS OF PROJECT AWARD FOR NAVY MILITARY CONSTRUCTION

Robert J. Thompson
Lieutenant Commander, United States Navy
BSCE, University of Colorado at Boulder, 2009

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2020

Approved by: Lyn R. Whitaker
Advisor

Samuel E. Buttrey
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The current process that Naval Facilities Engineering Command (NAVFAC) uses to bring Military Construction (MILCON) projects from concept to contract award fails to provide reliable and timely results. This poor performance leads to the untimely delivery of critical facilities, which directly impacts warfighting and power projection capabilities. It erodes the professional reputation of NAVFAC, leaving supported commanders to question whether an essential product or service will arrive on time.

Today, there are many internal NAVFAC teams exploring process-improvement opportunities across the entire construction timeline, of which pre-award is only one piece. However, many of these efforts are limited in terms of number of projects or project factors considered.

The focus of this thesis is to analyze projects from initial documentation up to contract award. To accomplish this, we linked two widely used but independent databases to capture the complete documented life cycle of hundreds of Navy-executed MILCONs. Once linked, dozens of project factors were collected, analyzed, and used in the development of various machine-learning models to assess their influence on project award performance.

Our analysis highlights several factors among existing and newly constructed project metrics that appear to greatly influence the timeliness of project award. This collection of potentially high-influence factors can then help NAVFAC further focus its ongoing process improvements.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

BCI	Building Cost Index
BY	Budget Year
CSV	Comma Separated Value
CV	Coefficient of Variation
DB	Design Build
DBB	Design Bid Build
DR	Design Release
DS	Design Start
EPG	Electronic Project Generator
FSDDA	Final Solicitation Document Design Authority
FY	Fiscal Year
HQ	Headquarters
IT	Information Technology
K-M	Kaplan-Meier
MILCON	Military Construction
NAVFAC	Naval Facilities Engineering Command
P&S	Products and Services
PDA	Preliminary Design Authority
POM	Program Objective Memorandum
RFP	Request for Proposal
ROC	Receiver Operating Characteristics
SME	Subject Matter Expert
UIC	Unit Identification Code
UOM	Unit of Measure

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

In recent years, the United States has seen its global military superiority challenged at an increasingly rapid pace. New and emerging threats coupled with the speed of innovation have meant that military leaders are now seeking ways to increase the agility of war-fighting forces and the delivery of support on which they depend. Because of its prominent role executing the Navy's multibillion dollar Military Construction (MILCON) program, the Naval Facilities Engineering Command (NAVFAC) is actively pursuing process improvement initiatives aimed at increasing its agility in delivering essential products and services. NAVFAC defines the focus of its agility-centric efforts in its Strategic Design 2.0 released in 2019 (Naval Facilities Engineering Command 2019). Within this framework, NAVFAC challenges itself to pursue ways of reducing the delivery time of MILCON projects, which would directly enhance war-fighting capabilities by delivering essential facilities and infrastructure when and where they are needed most. Our research directly supports this effort by focusing on the identification of MILCON project factors that are most likely to influence the likelihood of a project awarding on time. We accomplish this by developing explanatory predictive models of award performance for 400+ MILCON projects using 52 project factors spanning the entire life-cycle of each individual project, from concept to award.

Our study relies upon the marriage of two independent NAVFAC web-based systems. First, we capture 10,761 MILCON project records from the Electronic Project Generator (EPG) system dating back to its initial implementation in the early 2000s. This database serves as a repository for project documentation prior to congressional approval/funding, and so is a rich resource for understanding a project while it is underdevelopment. Second, we collect details for 628 Navy MILCON projects from the eProjects database, representing a majority of all fiscal year (FY) 10–20 projects. eProjects is a work management tool which tracks project performance from funding through project award. When paired with EPG, it captures all pre-award project details. Our research finds that the two databases have never been merged on a large scale before, which means analysts often assess MILCON project performance without access to all available project

details. Therefore, in this thesis, we spend considerable time documenting how to link the databases and in describing the follow-on analysis of the combined data.

In this study, we focus our efforts on understanding contributing factors to the timeliness of project award as measured in two different ways. The first timeliness project performance metric is whether a project "awards" (that is, is awarded) within its assigned budget year (BY). This is important since an award outside of the project's BY requires a report to Congress and has lasting repercussions. The second measure of timeliness is whether a project awards before its assigned Headquarter (HQ) Execution Lock Date, which further defines when a particular project is needed or promised to be awarded. Since our combined dataset contains projects which had not awarded at the time of data collection, relative to one or both of our measures, we use survival analysis to inform our data sub-setting strategy. Using Kaplan-Meier estimators (Kaplan and Meier 1958) we quantify the bias that results from removing un-awarded projects from the dataset. We find that sub-setting the projects to include only those projects which have been awarded, does not unduly bias our analysis and so proceed with model development using only the 464 awarded projects.

The driver behind our predictive model development is not in fact prediction. We use predictive models to glean insights into what project factors appear most influential in determining a project's likeliness of awarding on time. Therefore, interpretability of the model is an important consideration in our final model selection. With this in mind, we fit several models using both classification trees and random forest algorithms. The classification trees are fit using the *rpart* package (Therneau et al. 2019) in R (R Core Team 2019) and pruned according to criteria for minimizing cross-validated error. The random forest models are fit using both the *cforest* function of the R package *party* (Hothorn et al. 2020) and the R package *randomForest* (Liaw and Wiener 2002). Because of the limited size of our dataset, we do not select separate training and test sets. Instead, we train and test all models on the entire dataset using 10-fold cross validation. After initial model development we iteratively improve model performance through a combination of predictor variable consolidation, modification, or deletion.

We find that that for both response variables the classification trees outperform the random forest models. These models do not establish causal relationships between the predictor and response variables, but identify project factors to focus future exploratory analysis. Table 1 provides a summary of the most influential project factors for predicting the timeliness of project award relative to each response variable of interest.

Table 1. Summary of High Influence Factors

Top Influential Variables Across Models		
Variable Importance Rank Order	Award in Project BY	Award Before HQ Execution Lock Date
1	**Region/Execution Team**	**Region/Execution Team**
2	**Acquisition Tool**	Variability of Total Project Cost (during development)
3	**EPG Record Count**	DoD Explosives Safety Board / Air Instal. Comp. Use Zone / Airfield / Electromagnetic Radiation / Wetlands
4	Total Project Cost (FY20 $)	**EPG Record Count**
5	Contaminated Soil or Water	Elapsed time: Preliminary Design Authority to Final Solicitation Document Design Authority (days)
6	Month Design Release Received	Variability of Category Code Unit Cost
7	Acquisition Method	Design Level
8	Known Contaminant	**Acquisition Tool**
9	Memo of Negative Decision	Authority to Advertise (month)
10	National Capital Region Approved	Preliminary Design Authority (years in advance)

References

Hothorn T, Hornik K, Strobl C, Zeileis A (2020) party: A computational toolbox for recursive partitioning, R package version 1.3-5. https://cran.r-project.org/web/packages/party/party.pdf

Kaplan E, Meier P (1958) Nonparametric estimation from incomplete observations. *J. of the Am. Stat. Association* 53(282), https://doi.org/10.2307/2281868.

Liaw A, Wiener M (2002) randomForest: Breiman and Cutler's random forests for classification and regression, R package version 4.6-14. https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Naval Facilities Engineering Command (2019) External summary memo. Strategic Design 2.0. Washington, DC. https://hub.navfac.navy.mil/webcenter/portal/Strategic_Design.

R Core Team (2019). R: A language and environment for statistical computing, version 3.6.1. https://www.R-project.org/.

Therneau T, Atkinson B, Ripley B (2019) rpart: Recursive partitioning for classification, regression and survival trees, R package version 4.1-15. https://cran.r-project.org/web/packages/rpart/rpart.pdf

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

This thesis would never have happened without the mentorship, expertise, and patience of Drs. Lyn Whitaker and Samuel Buttrey. Thank you!

I would also like to express my thanks to CAPT Gordon (Tres) Meek III and Mr. Karthik Bharat for their tireless support and direction. Thank you!

As with everything worth doing in my life, I would like to thank my Thompson girls for their love, support, humor, and understanding. Thank you!!!!!

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

In 2018, the Chief of Naval Operations (CNO) released his Strategic Design 2.0 in which he calls for increased agility across the full spectrum of warfare capabilities (Department of the Navy [DON] 2018). The use of the term agility applies to more than just the direct operational warfighting capabilities of our fighting force. Agility also speaks to the shore facilities and infrastructure that warfighters rely upon to conduct their missions. This means that the Naval Facilities Engineering Command (NAVFAC) has a prominent role to play in order to drive innovation and process improvement initiatives. As the Navy's executing agent for Military Construction (MILCON), NAVFAC directly manages the life-cycle technical oversight of a program that is budgeted for $2.1 billion dollars in fiscal year (FY) 21 alone (Navy Office of Information 2020). Besides the monetary import of this program, each individual MILCON project is a large construction effort which requires a substantial amount of time to conceive, develop, design, and execute. In fact, the life cycle of a project from defined need to project completion is many years. Therefore, even modest process improvements which shorten the timeline associated with MILCON project delivery would significantly increase NAVFAC's agility in supporting warfighting needs. The focus of this thesis is to identify, using data analysis and machine learning, project features that influence the early part of a project's timeline. Our goal in doing so is to identify for NAVFAC high-impact areas on which to focus future analysis efforts.

In response to the CNO's directive and in appreciation of the prominent role NAVFAC's products and services (P&S) play, the Navy's Chief of Civil Engineers issued a NAVFAC specific Strategic Design 2.0 directive (NAVFAC 2019c). The directive reiterates the need to increase agility and highlights NAVFAC-centric focus areas for improvement across its many programs, of which MILCON is one. One specific focus area calls for the need to produce P&S in a timely manner so that they are still relevant once delivered (NAVFAC 2019c). This focus area establishes the goal of reducing the current design phase of Navy MILCONs to under one year. This is not a trivial undertaking

considering that an initial analysis performed by the working group established to implement this goal finds that the current design phase takes over two years on average (Meek 2019a). For the purposes of the one-year goal, the design phase is defined as beginning once the project has received congressional funding and ending once the construction contract has been awarded. However, there is also a significant period of time and work effort associated with getting a project documented, developed, and prioritized for congressional funding. Work performed during this pre-funding timeline is foundational for project scope setting and defining. Therefore, in order to truly understand a project's pre-award timeline a project tracked from the first time it is formally defined on a DD 1391 provides a more complete picture of performance. The significance of the DD 1391 is that it serves as the official project documentation for MILCONs and eventually serves as the defining document for project scope and funding. Therefore, the first draft of a project DD 1391 is truly the starting point for every MILCON. Collecting data from the first DD 1391 through project award provides complete visibility into a project's pre-award life cycle, which enhances follow-on analysis.

B. PURPOSE OF RESEARCH

The purpose of this research is to support NAVFAC's Strategic Design 2.0 goal of reducing the MILCON design phase to one year by identifying high-risk project features which are likely to influence a project's likelihood of awarding on time. In doing so, our goal is to help NAVFAC smartly allocate its limited resources towards high-impact process improvement opportunities through data analysis and machine learning. We accomplish this by identifying, fitting, testing, and selecting sufficiently accurate machine learning models from which predictions about award performance can be made. We then use the features of the best performing models to identify high impact areas for NAVFAC to focus in future exploratory analysis efforts.

C. RELATED WORK

Identifying and managing the project factors associated with cost overruns and time delays is a construction industry-wide question, which impacts both military and civilian construction alike. However, military construction does have unique challenges,

motivations, and priorities and so we consider both NAVFAC-centric literature in addition to civilian work.

1. Types of Project Influencers and Impacts

Larsen, Shen, Lindhard, and Brunoe (2015) seek to understand which specific project factors pertaining to commercial construction projects are most likely to have a negative impact on the performance elements of time, cost, and quality. Relying on past research and adding elements regarding construction quality, Larsen et al. (2015) conduct a series of hypothesis tests to determine if the negative influence factors impact all three of the performance elements in the same way. Larsen et al. (2015) determine that the negative influence factors impact the three performance elements quite differently. This is an important consideration since during the course of our research we focus exclusively on the question of time impact and not the other two, and so any conclusions we draw cannot be generalized to cost and quality without further research.

2. Machine Learning in Construction Delay Analysis

Attempting to understand risk factors associated with cost and schedule growth for construction activities is it not a new field of inquiry. However, with the emergence of machine learning methods and techniques, exploring these risk factors with machine learning is relatively new. Two recent studies of particular interest apply machine learning to understand risk factors associated with time delays for commercial construction.

In the first article, Yaseen et al. (2020) compare and contrast the performance of standard random forests and a random forest with genetic algorithm optimization. Yaseen et al. (2020) use data from 40 commercial construction projects and input from 300 industry experts to develop their pool of model parameters. Yaseen et al. (2020) find an increase in performance by pairing the genetic algorithm optimization with the random forest classifier. Additionally, Yaseen et al. (2020) provide a summary of the most likely project risk factors as a function of total project delay for their dataset. Some of the project risk factors they consider are those from the sources like the designer, equipment, contractor, labor, material, as well as others.

In the second study, Gondia et al. (2020) explore the use of decision trees and naïve Bayesian classifiers in order to determine which machine learning tool yields the best model for predicting project delay. Gondia et al. (2020) establish various performance indices and then use 10-fold cross-validation techniques to determine performance relative to these indices, as well as the overall relative performance of the models. Gondia et al. (2020) find that a naïve Bayesian classifier outperforms a decision tree by 3.9% when evaluating overall performance on their dataset. Of particular interest is how Gondia et al. (2020) fit these models, using a combination of risk source coupled with likelihood of risk occurrence and then weight these factors based on a severity index.

3. NAVFAC Specific Work

In addition to work performed by civilian researchers, NAVFAC is continuously conducting its own internal research to better understand and mitigate risk associated with the delivery of their P&S. To coordinate these efforts NAVFAC's Engineering and Expeditionary Warfare Center has been charged with collecting and disseminating academic research conducted by NAVFAC personnel. From this collection, Mack (2020) is a particularly interesting study since it parallels many of the elements in our own research. Mack (2020) conducts a statistical analysis of over 6,600 project records to determine the project features that most impact schedule delays. Additionally, Mack (2020) developed a machine learning model to identify project factors which influence project cost and timeliness of product delivery. The models Mack (2020) develops do not have a high predictive power and only use a small number of factors but still offer interesting observations regarding the relative importance of the project factors considered. Of note is that design level, design agent, and project cost are all important factors regarding project delays. Additionally, to a lesser extent, both the project's region and execution team are also important factors.

D. ORGANIZATION

In this thesis we first focus on the collection, manipulation, and analysis of available MILCON data. In Chapter II we introduce, and describe the collection and combining of data from two separate databases, the Electronic Project Generator (EPG) and eProjects,

on a scale not attempted before. The combined dataset then allows us to explore project performance by leveraging all data from the entire project's life cycle. In Chapter III we explore the combined dataset through univariate and multivariate analysis. Using these results, in Chapter IV, we fit numerous machine learning models to the data in order to predict award performance relative to the timelines of project award. Finally, conclusions and recommendations are given in Chapter V.

THIS PAGE INTENTIONALLY LEFT BLANK

II. DATA INTRODUCTION

In this chapter we discuss the potential sources of Navy MILCON-related data, identify the particular datasets used during the course of this research effort and explain how the datasets are merged to form a single unified dataset for follow-on analysis. The chapter then concludes with an initial exploration of the consolidated dataset and the resulting collection of project factors, which serve as the foundational dataset for this research.

A. AVAILABLE DATASETS

In the performance of its duties as the executing agent for Navy MILCONs, NAVFAC uses numerous information technology (IT) systems to track and manage project data from initial concept through P&S delivery. A collection of these systems represented by yellow boxes can be seen in Figure 1, which is captured from a continuous learning module provided to NAVFAC personnel. A detailed description of these systems can be found in Brown et al. (2020).

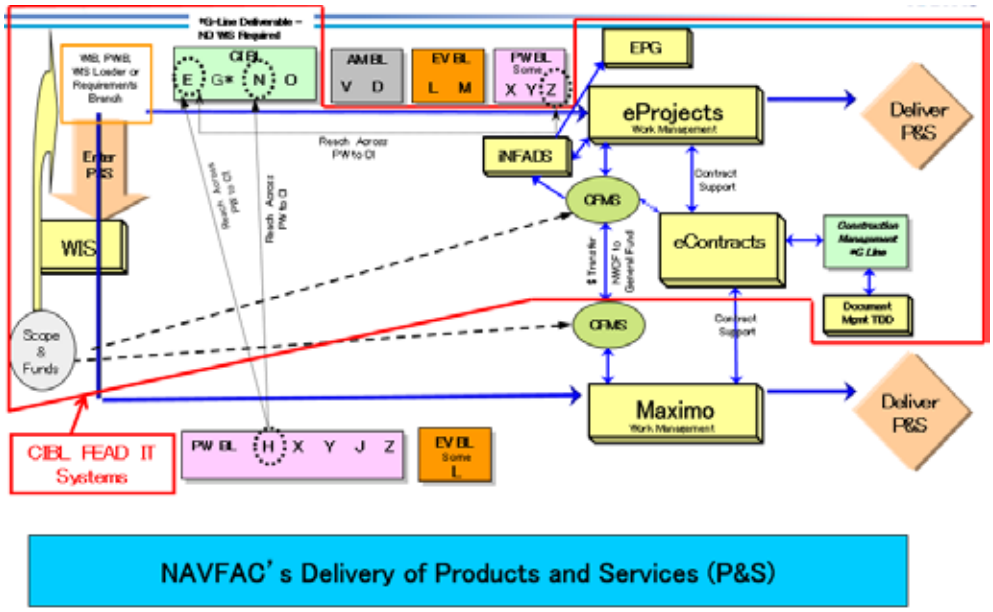

Figure 1. NAVFAC IT systems. Source: Brown et al. (2020).

As expected there are redundancies, overlap, and transitional points across these databases that are not always transparent to the end user. In this thesis we focus on project data from initial concept up through contract award. This equates to a collection of data spanning the EPG and eProjects databases. Although both systems contain information from the same database, EPG and eProjects act independently from each other. EPG serves as a repository for project documentation in the form of a project DD 1391 in various stages of development and tracks the project from initial concept through congressional funding. This then serves as a transition point from EPG to eProjects, where eProjects then acts as a work management system from congressional funding received through contract award. Therefore, an essential task of this research is to identify and leverage a way of linking the information contained in these two systems together for a distinct set of projects.

1. Electronic Project Generator Database

The EPG database serves as a repository for all project DD 1391s. Each DD 1391 captures detailed project information, costs, and so on, and is used to request project funding. As a project's DD 1391 progresses from initial generation all the way to enactment by Congress (if the project makes it that far), the DD 1391 is tracked through the creation of various records in EPG corresponding to status level changes or content changes and updates. Projects are no longer tracked in EPG after Congressional approval and funding, identified by a closeout record in EPG with status level "As Enacted." Therefore, EPG is an invaluable resource from which to gather detailed project description/details, associated costs, and to track cost and scope changes during the course of project development.

EPG stores the information from each DD 1391 in a series of inter-connected tables. In some cases, individual blocks within the DD 1391 are partitioned into several independent tables. The particular information being sought from EPG has to be clearly articulated so that the correct collection of tables can be extracted. This is a particular challenge of exploring this database and is discussed in more detail in Chapter V. The data pulled from EPG takes the form of six comma separated value (CSV) files (Bryce, 2019c), one for each of the six tables we consider from EPG. These tables include data from all available records (38,495 unique projects) but do not include all possible data fields. We

focus on collecting the principal project details comprising information found in Blocks 1–9 & 12 of the DD 1391s. Figure 2 shows the database schema developed in coordination with the database manager (Bryce, 2019b). The line endings for each connector in the schema indicate the type of variable relationship between tables. A split ending indicates "many" and a cross ending indicates "one." For example, a connecting line featuring both of the two possible line endings indicates a "one to many" variable relationship. Each table features numerous fields but for brevity not all table fields are included in Figure 2.

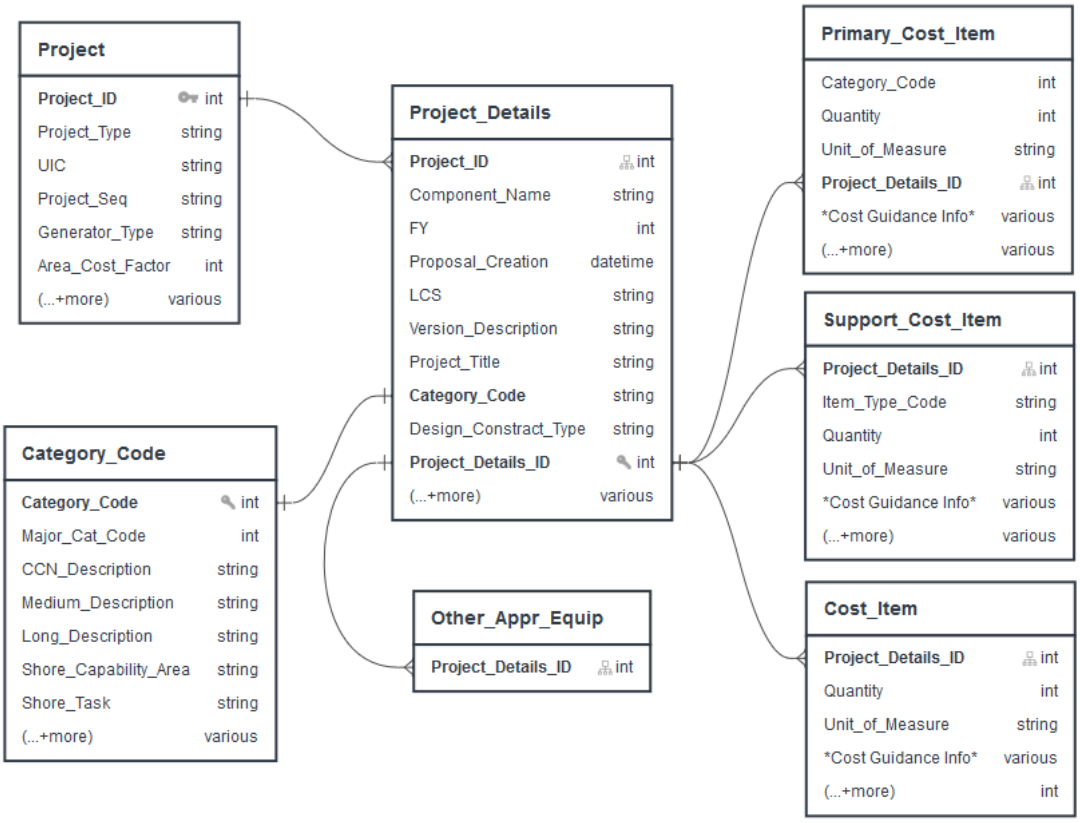

Figure 2. Schema for EPG Data Extraction and Consolidation

As a starting point, we leverage the schema from Figure 2 to pull the data into the statistical computing environment R (R Core Team 2019) for follow-on analysis and manipulation. The tables from EPG use a numeric label referred to as the "Project_Id" code to identify individual projects. This code is meaningless outside of the context of the

database architecture but is used to track distinct projects. Figure 3 gives a snapshot of total number of unique records organized by project type.

Figure 3. Breakdown of EPG Records by Project Type

Most project records correspond to Special Projects. However, this study focuses only on MILCONs. There are 10,761 MILCON records out of 38,495 project records dating from the early 2000s. As a first step, the "Project_Ids" corresponding to MILCONs are identified and used to extract the set of MILCON projects.

Though the "Project_Ids" are unique, the database may have several (possibly dozens) associated records for any project identified by an individual record's "Project_Details_Id." Each time any modification or change occurs to a project's DD 1391, a new record is created, a "Project_Details_Id" assigned, and the EPG status changed if appropriate. Therefore, after filtering the "Project_Ids" for only those corresponding to MILCON projects, we develop a method to capture only those EPG records that are relevant. This method is developed in coordination with NAVFAC Headquarter (HQ) Subject Matter Experts (SMEs) and applies the approval process associated with the DD 1391s. Since EPG captures various drafts, only the records corresponding with the various approval levels have accurate, leadership reviewed and approved information. We take this approval process in the form of the "EPG Status Levels" flowchart (Figure 4) from recent internal NAVFAC guidance (Bharat 2019).

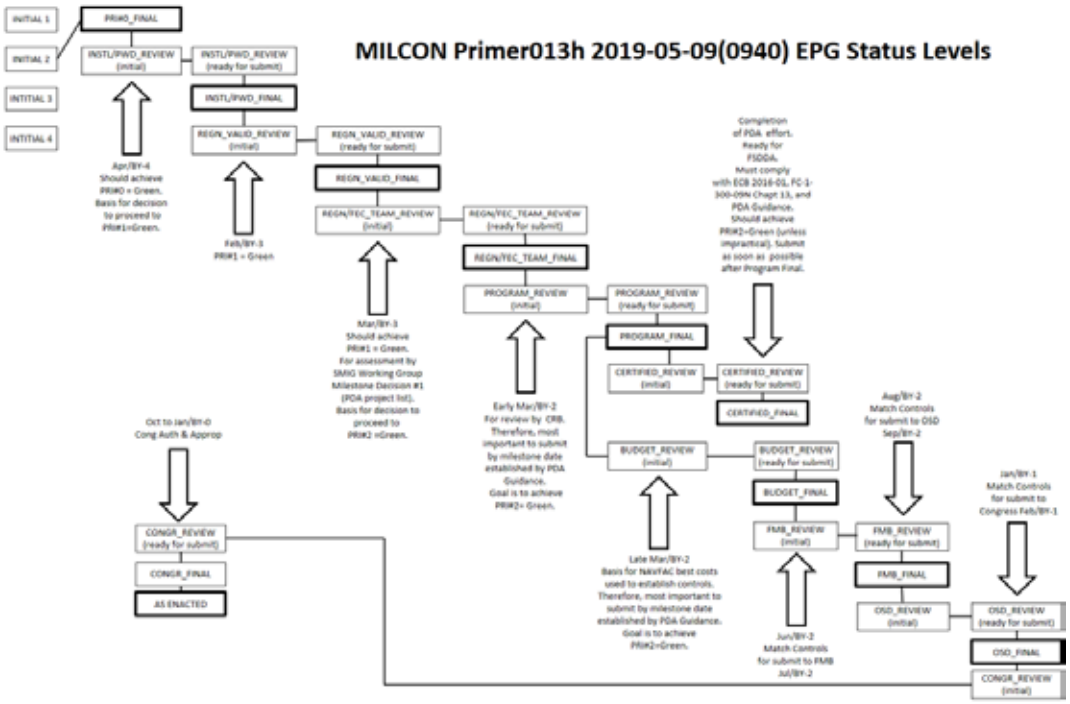

Figure 4. EPG Status Levels from MILCON Primer. Source: Bharat (2019).

However, since the EPG database and MILCON approval process has evolved over time, some status levels have undergone name changes. After reviewing all possible record status levels in EPG, we make the revisions shown in Table 1 to individual records to be sure all applicable project information is captured (Bryce 2019a).

Table 1. Consolidated EPG Status Levels

Available Status Levels	Revised Status Levels
ACTY_FINAL INSTL/PWD_FINAL	INSTL/PWD_FINAL
REG_VALID_FINAL	REG_VALID_FINAL
REGN/FEC_TEAM_FINAL REGIN_FINAL	REGN/FEC_TEAM_FINAL
FEC_FINAL PROGRAM_FINAL	PROGRAM_FINAL
CERTIFIED_FINAL	CERTIFIED_FINAL
NAVFACHQ_FINAL BUDGET_FINAL	BUDGET_FINAL
FMB_FINAL	FMB_FINAL
OSD_FINAL	OSD_FINAL
AS ENACTED	AS ENACTED

After applying the required status level changes, the MILCON records are then further filtered by retaining only records marked as having a "Final" version of one of the nine key milestones in the approval flowchart. Figure 5 illustrates the breakdown of the number of MILCON project records by record status level.

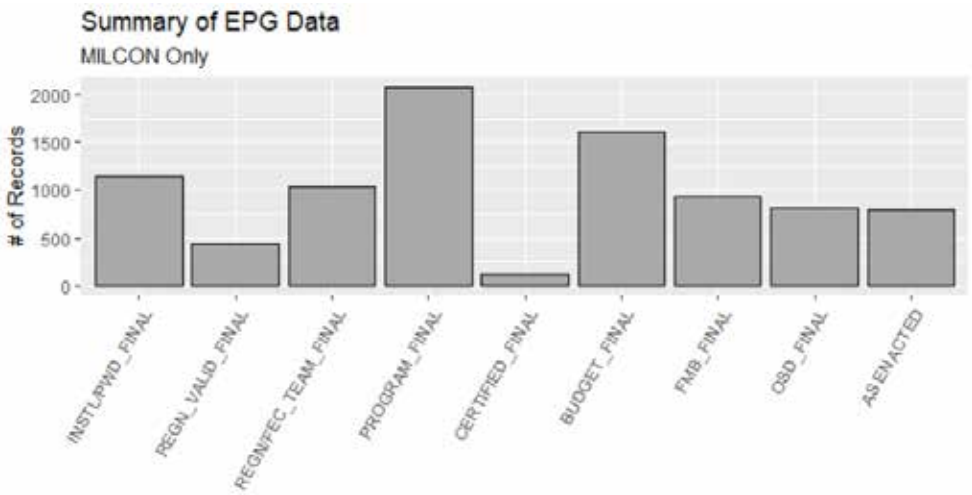

Figure 5. Number of MILCON Records by EPG Status Levels

2. eProjects Database

eProjects is an element of the larger NAVFAC suite of applications known as ieFACMAN and is used as a work management system by NAVFAC's Asset Management business line. Besides containing some project details, eProjects is used to track a project's progression through various milestones from the date that Preliminary Design Authority (PDA) is received to project award.

Our research focuses on an eProjects dataset for FY10-20 MILCONs (Meek 2019b). The dataset contains 648 individual records for 628 unique projects. To prevent an unintentional loss of data, we reconcile the duplicated records prior to linking eProjects to EPG. In most cases duplicates result from a placeholder Program Objective Memorandum (POM) DD1391 submission and therefore these duplicates are removed from the dataset. One project (N60514-P343) has a duplicate record with notes showing that NAVFAC canceled the project, and so we also remove it from the dataset.

After the removal of duplicate entries, the dataset is grouped by MILCON dollar amount, region, and by design level (Figure 6). The possible MILCON design levels are Design Bid Build (DBB) and Design Build (DB). The principal distinction is that for DB projects the design is the responsibility of the same entity performing the construction. For a DBB project the design and construction are performed by separate entities, in a linear workflow. We identify these groupings as having specific interest to the ongoing NAVFAC analysis efforts. One project has a region listed as N/A or missing. This project's Unit Identification Code (UIC) is NC1002 and the assigned P# is 923.We assign the region for this project to be JK Marianas, which appears to be the best fit.

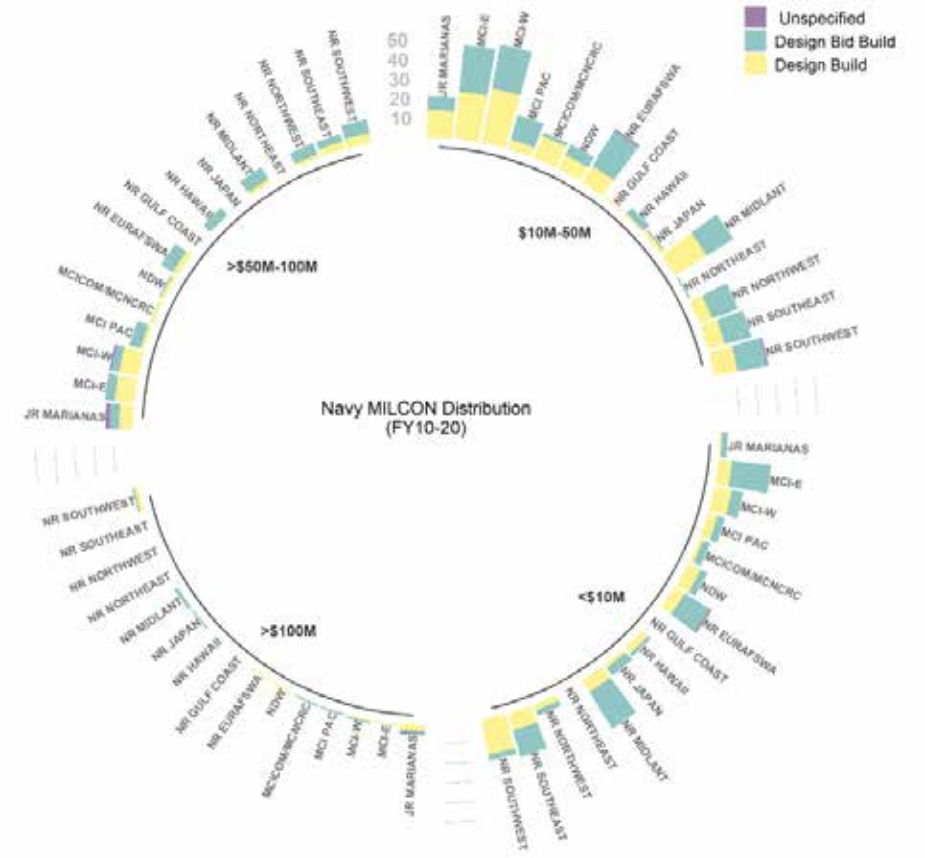

Figure 6. Navy MILCON Distribution by Count (FY10–20)

Figure 6 illustrates the frequency of MILCONs as the project cost increases, starting in the lower right-hand quadrant and moving counter-clockwise. Most projects fall below

a total project cost of $50M with very few projects over $100M. Taken as a whole, the total number of DBB projects is approximately equal to the total number of DB projects, though some regions appear to prefer one over the other depending on project dollar amount.

B. LINKING RELEVANT DATASETS

This section describes how the two relevant databases are merged, presents an initial analysis of the joint dataset, and defines the final dataset used for all follow-on analysis conducted as part of this research effort.

1. Linking EPG and eProjects (creation of a combined dataset)

Joining EPG records with their corresponding projects tracked in eProjects makes all the project metadata immediately available for analysis. However, these two systems are not connected. To join them, the schema developed and presented in Figure 7 is used to link each eProjects record to its corresponding "Project_Id" in EPG.

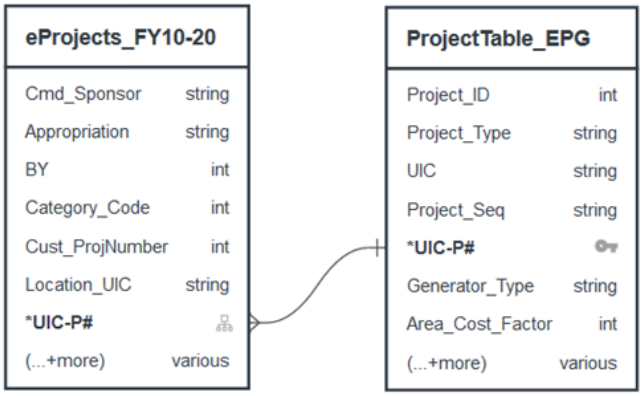

Figure 7. Schema to link EPG to eProjects dataset

This schema requires the creation of a unique "UIC- P#" label within each table independently. It is important to note that each database is sensitive to leading zeroes in their respective "P#" data fields, and therefore the zeroes must carry over faithfully into the newly created "UIC-P#" label. Without the leading zeroes it is possible to have several non-unique "UIC-P#" labels, resulting in an incorrect matching of records across eProjects

and EPG. Using this labelling technique, a unique label is generated for 10,731 MILCON projects of the possible 10,756 projects in EPG. In 24 cases the "UIC-P#" label corresponds to two or more "Project_Ids" but does not affect any project corresponding to the eProjects dataset.

The EPG database is then subsetted by filtering for only "Project_Ids" corresponding to records with a "UIC-P#" label that match one from the eProjects dataset. This results in a subset of EPG data for 622 of the possible 628 unique projects from the eProjects data. Table 2 shows the projects which appear in eProjects but do not have an EPG record that corresponds with the "UIC-P#" number pairing.

Table 2. Projects from Dataset without Corresponding EPG Record

UIC-P# Pairs					
N33191-312	N62863-2001	N63005-960	M00318-86167	M67001-678x	N83447-136

We inspect the pairs presented in Table 2 to determine whether there is an obvious explanation as to why these projects might appear within eProjects but not EPG. Since project data is often input by hand, it seems reasonable to explore obvious miss-type errors in the project "P#'s," resulting in an unmatched "UIC-P#" label.

By inspection, the project details for M67001-678x match exactly the details for the EPG record with label M67001-678. Therefore, we change the "UIC-P#" label in the eProjects data from M67001-678x to M67001-678 and then pair it with its corresponding EPG record. This project only has a corresponding "Initial" record in EPG. So even though it is added back into the dataset at this stage, it is further explored in the next section for possible exclusion.

For this thesis, we consider projects complete and useable for analysis only if they have (at a minimum) an eProjects and EPG record. Since we find no obvious explanation for why the remaining five projects do not have a corresponding EPG record, we exclude them from the dataset at this stage.

2. Exploration of EPG/eProjects Combined Dataset

The EPG/eProjects combined dataset acts as the upper limit in terms of dataset scope for this thesis. Using this dataset as a starting point, we conduct an initial analysis to determine whether it should be further reduced because of limited or incomplete project detail quality across the combined records. The following are some noteworthy observations.

a. Absence of "Final" EPG records (any level)

Out of the 623 projects, 11 of them do not have a corresponding "Final" EPG record at any status level. Those 11 projects only have "Initial" records within EPG. Those projects are identified in Table 3.

Table 3. Projects with "Initial" EPG Records Only

UIC - P#	EPG PROJECT_ID	UIC - P#	EPG PROJECT_ID
N60042-116	33877	N60508-278	144239
N62863-708	45565	N61065-240	85805
N63005-014	58023	N61755-630	102789
M67001-678	50463	N47609-003	102367
N32411-056	52043	N62863-733	134751
M67400-305	101986		

The presence of only an "Initial" record means that EPG cannot be used to analyze metrics like cost growth, labor effort required, etc., for these projects. However, the "Initial" records will still contain amplifying project details that can enhance the analysis of the project details found in the eProjects dataset, most notably questions regarding schedule growth. Therefore, we inspect each record for completeness in the corresponding eProjects project entry. Of the 11 projects, 10 of them have significant detail in their corresponding eProjects record and therefore we keep them in the dataset. The one project without substantial detail is N60042-116, and so we remove this project from the dataset at this point.

b. Projects without "As Enacted" records

There are only 520 unique "As Enacted" records that correspond to the eProjects dataset. That means, taking into consideration the 11 projects mentioned above, there are 92 additional projects from the eProjects dataset that never reach the last EPG stage of "As Enacted." For follow-on analysis it is helpful to know where in the EPG routing process each project resides to determine how complete and useable the data is.

Figure 8 shows the furthest step each project DD 1391 has reached within the EPG routing process.

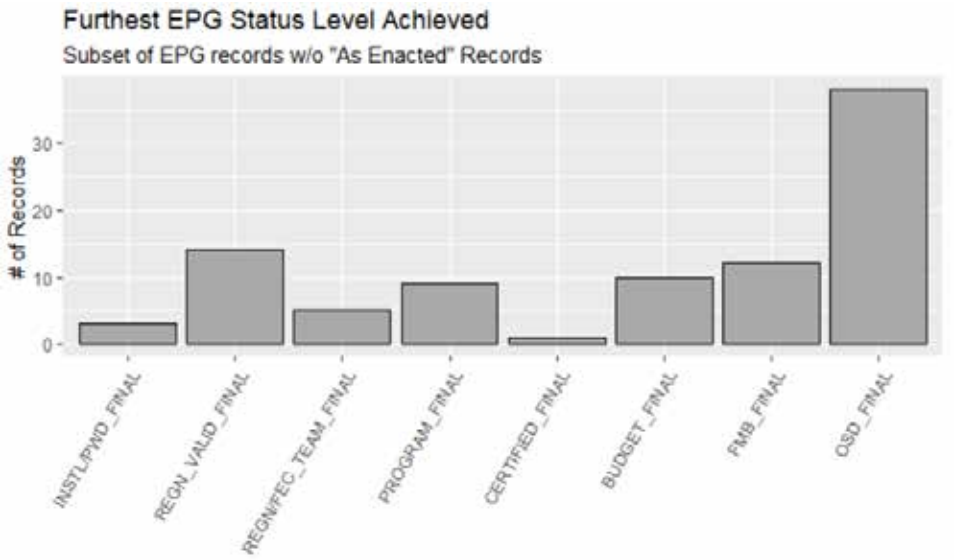

Figure 8. Furthest EPG Status Level Achieved

Figure 8 is generated by first isolating the "Project_Ids" corresponding to the projects without an "As Enacted" record in EPG. Then the EPG status levels are considered in reverse order. At each level the "Project_Ids" are captured for projects, having a record matching the current status level. For each successive step the "Project_Ids" are kept if they have not already been for a later stage. The result then illustrates the single furthest status level achieved by each project.

From Figure 8 we can see that most projects that have not reached the "As Enacted" step at the time of the data extraction are most often in the later stages of the process. The

single most common stage is the “OSD_Final,” which is the immediate predecessor to the “As Enacted” step. Therefore, the corresponding information in EPG will be assumed to be reasonably well vetted and complete.

C. FINAL PROBLEM DATASET DEFINED

As discussed previously, the dataset used for this thesis spans two NAVFAC databases. The first, EPG, captures detailed MILCON project information including scope, cost information, etc. The second, eProjects, serves as the work management system for delivering MILCONs to the end user and captures some project details and scheduling information. The question this research addresses is, “what factors affect or predict a project’s likelihood of awarding on time?” Therefore, combining the data from the two databases enhances this analysis by making all of a projects metadata available as factors in a predictive model.

Merging the two datasets together is constrained by the most restrictive subset, which in this case is the FY10-20 eProjects dataset. This dataset has 648 project entries, but we do not use all of them. Below, we summarize how and why the eProjects dataset is further reduced based upon the analysis of the previous sections:

- 20 eProject entries are found to be duplicates of otherwise reported projects
- Five eProject entries do not have a corresponding EPG record
- One project (N60042-116) lacks significant detail in both EPG and eProjects

The resulting dataset after these deletions has 622 Navy MILCON projects between FY10 and FY20.

From the initial exploration of the combined dataset the following are important considerations for handling the final problem dataset moving forward:

- The way that project documentation is routed within EPG is not consistent. In other words, not every project will have an EPG record at every status level. Entire steps may have been skipped.

- Not all records have a "Final" record at any level. Important project information can be found within the "Initial" records but certain variables will not be available for these projects.

- Not all records have fully completed the EPG routing process culminating in an "As Enacted" record

D. METHODOLOGY

In the previous section we define the dataset for this thesis. This section discusses the scope of response and predictor variables we use in our research. Our focus is on the timeliness of project award and so we first start by defining how timeliness is measured in the form of our chosen response variables. Next, we outline the collection of predictor variables from the combined EPG/eProjects dataset and provide brief variable descriptions. Finally, we highlight the limitations and assumptions of our research.

1. Response Variables

There are two response variables of interest, both of which describe the timeliness of a project award. The first is the actual award relative to the end of the project's budget year (BY). This metric is a Congressional reporting requirement. The second is the actual project award relative to the project's Headquarter (HQ) Execution Lock Date. This is a way for NAVFAC leadership to track the date that the MILCON contract award is required by.

We can report both responses as simple binary variables showing whether the project awarded on time (Yes/No). On time means that the project awarded either before the end of the project's BY or before the listed HQ Execution Lock Date. The responses can also be measured and reported as a number showing how many days in advance, or how late, the project awarded relative to either the end of the project's BY or the HQ

Execution Lock Date. During this research effort, we explore each response in both a binary and numeric format.

2. Scope of Combined Database Project Factors

This section provides three tables (Tables 4, 5, and 6) with descriptions of predictor variables by variable type used in our analysis. In general, we carry over the variable format from the parent database unless it makes sense to do otherwise. For a complete listing of all variables, including from which database they are captured from, see the appendix.

a. Numeric Variables

Table 4. Numeric Variables

Variable	Description
PDA	The date that the Preliminary Design Authority (PDA) was received. This provides the authority to obligate and expend MILCON design appropriations up to submission of the Certified Final DD 1391
FSDDA	The date that the Final Solicitation Document Design Authority was received. This provides authority to obligate and expend MILCON design appropriations for design efforts up to construction contract award
PDA to FSDDA (days)	Total elapsed time between the day that PDA and FSDDA were issued
DS Actual	The actual date that the project design was started
FSDDA to DS (days)	Total elapsed time between when the Final Solicitation Document Design Authority was received and when the project design was actually started
ATA	The date the Authority to Advertise was received
DR Actual	The date the design was complete
DS to DR (days)	Total elapsed time between when the project design was started and when the design was complete
DR Lock	Expected date that the design was complete
RFP Actual	The date that the Request for Proposal (RFP) was issued
DR to RFP (days)	Time from when the design was complete to the day that the Request for Proposal was issued
RFP Lock	Expected date that the Request for Proposal was issued
AWD Actual	The date that the project contract was actually awarded

Variable	Description
RFP to Award (days)	Time from when the Request for Proposal was issued to when the contract award was made
Total Duration of Design Phase (days)	Total elapsed time from PDA to project award
HQ Execution Lock	Date that the contract award is expected and advertised to leadership
AWD to Execution Lock (days)	Time from project contract award to when the award was expected. Negative values indicate an early award (desirable) and positive values indicate late award (undesirable)
Proposal Creation Date	Date that the EPG record within the Project Details table was created
Project Maturity	Total time elapsed from the first EPG project record to the final "As Enacted" record
Project Dollar Amount (project BY $s)	Total project dollar amount
Standardized Project Dollar Amount (FY20 $s)	Total project dollar amount standardized to FY20 dollars. Escalation factor determined by using the NAVFAC Building Cost Index (BCI) and the following equation: Escalation factor = (FY 2020 Index)/(Project BY Index)
Production Cost (project BY $s)	Cost of producing project plans and specifications
Other Design Cost (project BY $s)	Other costs associated with developing a project design not covered by the Production Cost
Contract Cost (project BY $s)	Cost of design elements satisfied by contract action
In House Cost (project BY $s)	Cost of design elements satisfied by internal NAVFAC work effort
Total Design Cost	Total Design Cost = Production Cost + Other Design Cost -OR- Contract Cost + In House Cost
Cost Growth Thru Development (FY20 $s)	Difference in standardized project dollar amount from the first project record to the last
Proposed ACF	Adjustment factor to account for the impact of project location on cost

Variable	Description
Size Adjustment Factor	Adjustment factor to scale guidance costs to reflect true magnitude of work element
Original Unit Cost (project BY $s)	Element cost prior to adjustment and escalation
Unit Cost (project BY $s)	Cost after adjustment and escalation
Proposed Escalation Index	Proposed and applied escalation factor
Guidance Escalation Index	Suggested escalation factor

b. Categorical Variables

Table 5. Categorical Variables

Variable	Description	Levels	Level Description
Budget Year	Year of project funding and anticipated award	11	2010-2020
Category Code - Primary	Overarching category code for the entire project. Subjective determination based upon the principal project element	Multiple	-
Location UIC	Way of identifying intended project location	Multiple	-
UIC by Project Volume	Project locations- Grouped by project volume	Very High	>15 projects
		High	11-15 projects
		Medium	5-10 projects
		Low	<5 projects
Region	Geographic region of project	JR Marianas	Joint Region Marianas
		MCI-E	Marine Corps Installations Command - East

Variable	Description	Levels	Level Description
		MCI-W	Marine Corps Installations Command - West
		MCI Pac	Marine Corps Installations Command - Pacific
		MCICOM/MCNCRC	Marine Corps Installations Command
		NDW	Naval District Washington
		NR EURAFSWA	Navy Region Europe, Africa, and Southwest Asia
		NR Gulf Coast	Navy Region Gulf Coast
		NR Hawaii	Navy Region NR Hawaii
		NR Japan	Navy Region NR Japan
		NR MIDLANT	Navy Region NR Mid-Atlantic
		NR Northeast	Navy Region NR Northeast
		NR Southeast	Navy Region NR Southeast
		NR Southwest	Navy Region NR Southwest
Responsible Component	NAVFAC element that is responsible for the project	EUR	NAVFAC Europe and SW Asia
		FE	NAVFAC Far East
		HI	NAVFAC Hawaii
		MAR	NAVFAC Marianas
		ML	NAVFAC Mid-Atlantic
		NW	NAVFAC Northwest
		SE	NAVFAC Southeast
		SW	NAVFAC Southwest
		WASH	NAVFAC Washington
Responsible Team		High	>100 team projects

Variable	Description	Levels	Level Description
	NAVFAC team responsible for overall project - Grouped by project volume	Medium	50-100 team projects
		Low	<50 team projects
Execution Team	NAVFAC team responsible for project execution - Grouped by project volume	High	>50
		Medium	6-50
		Low	<5
Acquisition Tool	Type of acquisition tool used for project	Unspecified	Acquisition tool not indicated within dataset
		MAC	Multiple Award Contract
		Standalone	Single standalone contract
		Other	Various infrequently used acquisition tools
Acquisition Method	Acquisition method used for project	RFP - BVSS (One Step)	
		RFP - BVSS (Two Step)	
		RFP - LP	
		Other	
Current Status Level	Status level of particular EPG Project Details record	Initial	
		Instl/PWD_Final	
		Reg_Valid_Final	
		Regn/Fec_Team_Final	
		Program Final	
		Certified Final	
		Budget Final	
		FMB Final	
		OSD Final	
		As Enacted	
Number of EPG Status Levels Achieved	Total number of "Final" EPG Status Lvls the project has achieved	1-9	
Standard Design	Is the project design considered standard	Yes	
		No	
		Unspecified	

Variable	Description	Levels	Level Description
Project Dollar Amount Groups	Total project costs grouped by magnitude	<$10M	
		$10M-50M	
		>$50M-100M	
		>$100M	

c. Binary Variables

Table 6. Binary Variables

Variable	Description
Execution Team by Project Volume	High: Executing team has a high volume of reported projects Low: Executing team has a low volume of reported projects
Design Level	DB: Project executed as Design Build DBB: Project executed as Design Bid Build
Design Agent	IH: Design performed internal to NAVFAC AE: Design performed by a third-party architectural and engineering firm
Design / Construction by Other	Yes: Design or Construction executed by party other than NAVFAC No: Design and Construction executed by NAVFAC
Award in Planned Year	Yes: Project was awarded in the projects Budget Year No: The project was not awarded during its Budget Year
PE Used	Yes: A parametric estimate was used No: A parametric estimate was not use
ES LCA Performed	Yes: An ES LCA was performed for this project No: An ES LCA was not performed for this project
Active	Yes: EPG record is listed as Active. Typically indicating the most recent project record No: EPG record is Inactive, indicating a more current record exists
EPG Routing Complete	Yes: Project has a record with “As Enacted” status level No: Project does not have an “As Enacted” record, indicating the project is still being tracked/routed within EPG

3. Limitations and Assumptions

One limitation of our research is that we focus only on the analysis of MILCONs. We do not consider or explore any other project type during this study. We elect to do this because of the priority that NAVFAC leadership has placed on improving the MILCON program and because of the sheer magnitude of time savings possible with even minor process improvements to the program. There is a significant amount of data pertaining to Special Projects, within EPG and elsewhere, and so this is an opportunity for future exploration.

Another limitation is that our analysis is performed on a static dataset extracted from two non-static databases. Many project managers regularly update project details across both parent databases, and so our analysis suffers from data updates not captured at the time of data extraction. We assume that those changes are minor and would only affect a small percentage of projects explored. Significant changes are unlikely to affect the pre-award timeline of projects already awarded at the time of data collection, and so any changes would primarily affect un-awarded projects. Because of this and the response variables of interest, we later explore how to divorce the dataset from projects within the dataset but un-awarded by 16 January 2019.

A final limitation of our study is that we only assess award performance relative to timeliness of award. We do not explore elements pertaining to cost growth, though this is also another very important element of MILCON performance. We focus on time instead of cost because of the emphasis the NAVFAC Strategic Design 2.0 places on enhancing agility. Therefore, any conclusions drawn in this study pertaining to process improvements are only relevant in terms of time savings, not cost savings. We cannot assume that risk factors regarding the timeliness of project award are also risk factors regarding project cost (Larsen et al. 2015).

During this research effort, we make various assumptions regarding the accuracy of data from both databases. We consider the data to be equally valid across the two databases when the same project factor is reported. In the rare instance that the two databases differ, we either make an informed determination or contact a NAVFAC SME

to de-conflict them. We also assume that information reported in a later EPG status level record was more accurate than the same data reported in an earlier record.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DESCRIPTIVE STATISTICS

A. INTRODUCTION

In Chapter II we explain the relevant sources of MILCON project data, outlined the initial project dataset, and define the response variables. In this chapter, the data is further explored to gain insights into what project factors (or combinations of factors) are most likely to influence or disrupt a timely contract award relative to the end of the project's BY and/or HQ Execution Lock Date. Exploration of a timely award requires that an observed project be awarded on or prior to 16 January 2019 when the data was collected (Meek 2019b). Therefore, a portion of this chapter addresses this constraint.

This chapter begins by first discussing what portion of the MILCON projects captured within the dataset are eligible for use in developing the various descriptive statistics and modeling in our research. Besides defining eligibility, we show that omitting un-awarded projects does not unduly bias our analysis.

Next, we present the results of a univariate and multivariate analysis of some important relationships between various project details and timeliness of award. The specific relationships presented within this section are those highlighted by various SMEs as "likely" to have a discernable influence on project award (Meek 2019a).

Finally, we explore opportunities for enhancing the dataset by leveraging the contents of the EPG database. Specifically, variability in project packages (relative to cost and scope) during development and project document maturity are widely believed to influence a timely award and so we seek to study this belief.

B. REQUIRED DATA SUB-SETTING

As discussed previously, a timely award is defined as an award relative to the end of the project's BY, as well as, relative to the HQ Execution Lock Date. Both of these are considered important performance metrics and so we assess and discuss them in parallel. Since the dataset was collected on 16 January 2019 and represents a snapshot of the FY10-

20 Navy MILCONs, not all projects within the dataset have an award date. Figure 9 illustrates the number of un-awarded projects by their BY.

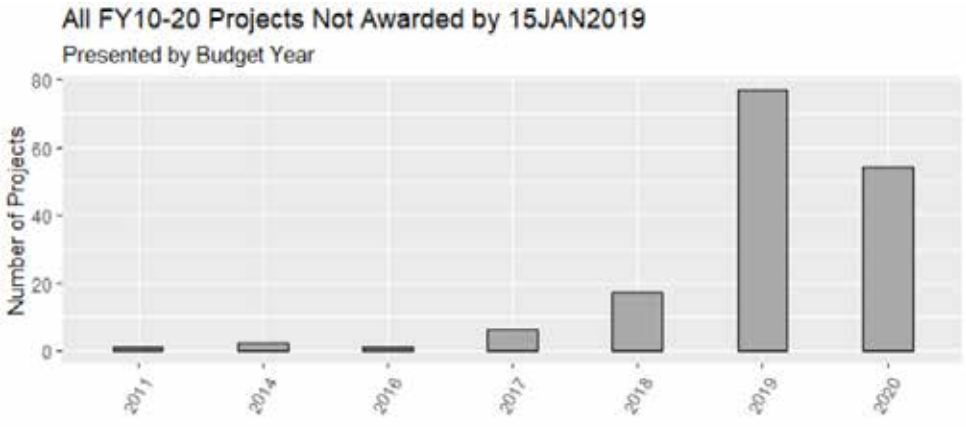

Figure 9. Distribution of Un-awarded Projects (FY10–20)

Since many projects are in the process of awarding there is the possibility that omitting them from the dataset could mean that results are biased in favor of projects with a shorter contract award timeline. In other words, the dataset consists of right-censored observations. Therefore, in order to determine the possible magnitude of biasing, a Kaplan–Meier (K-M) estimator (Kaplan and Meier 1958) is used to estimate the survival function for two distinct timelines: start of project BY to project award, and date of project Final Solicitation Document Design Authority (FSDDA) to project award. The survival function is the probability that the project is awarded after time *t*, where *t* is measured using one of the two timelines. For these calculations the *survival* package from R is used (Therneau et al. 2020). The following subsections discuss the resulting K-M curves and key findings.

1. Start of Project BY to Actual Award of Project

To fit the first K-M estimator, all FY20 projects are first removed from the dataset. We do this because the start of the BY for FY20 projects occurs after data collection. Then the start of each remaining project's BY serves as the starting point of the design timeline. We measure the time between this starting point and the project's actual award or date of

data collection, whichever comes first. Figure 10 shows the resulting K-M estimate of the survival function computed in three ways.

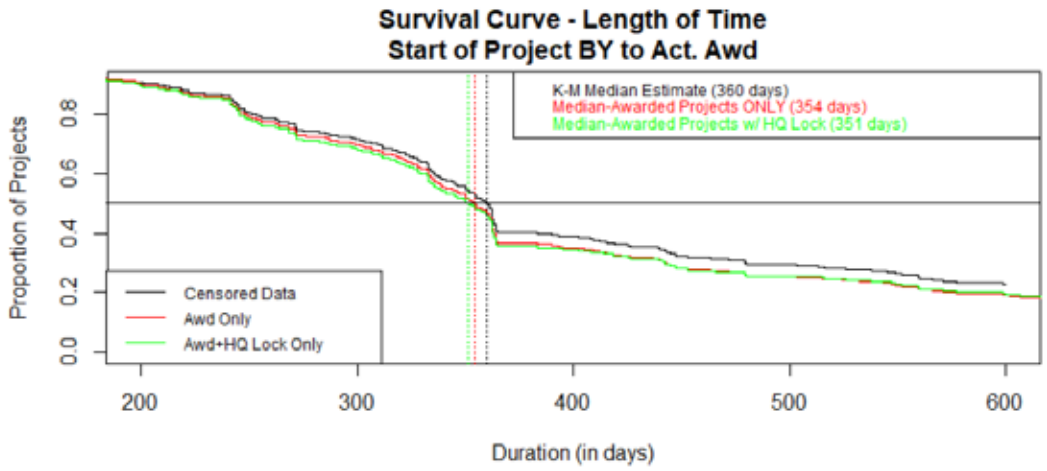

Figure 10. Survival Curve: Start of Project BY to Actual Award

Table 7 provides a summary of the median time to award as estimated from the censored data compared to the estimate derived from the "Awarded Projects" and "Awarded Projects with HQ Execution Lock Date" subsets (neither of which are censored).

Table 7. Results of Survival Curve: Start of Project BY to Actual Award

Number of Obs.	Number of Events	K-M Estimated Median (days)	95% Confidence		Awarded Projects Only		Awarded Projects w/ HQ Lock Date	
			LCL	UCL	Median (days)	Departure from K-M (%)	Median (days)	Departure from K-M (%)
568	464	360	351	362	354	-1.7%	351	-2.5%

It is clear that sub-setting the data and removing projects that have not yet awarded (i.e., the censored data) does not result in undue biasing of the results. Both of the subsets result in an estimated median design time length within the 95% confidence interval of the K-M estimated median based on the censored data and have only a marginal deviation as measured by percent deviation from the censored median. However, this particular timeline

does not take into account any FY20 projects. Therefore, we also define and explore a second design timeline.

2. Project FSDDA to Actual Award

The second K-M estimator is applied with the project design timeline measured by the day that a project received FSDDA to the day of project award or data collection. In this case the distribution of projects missing a FSDDA date is more spread out among the various years and so it is reasonable to again assess potential bias occurring as a result of sub-setting. Figure 11 and Table 8 give the resulting K-M estimates and summary statistics.

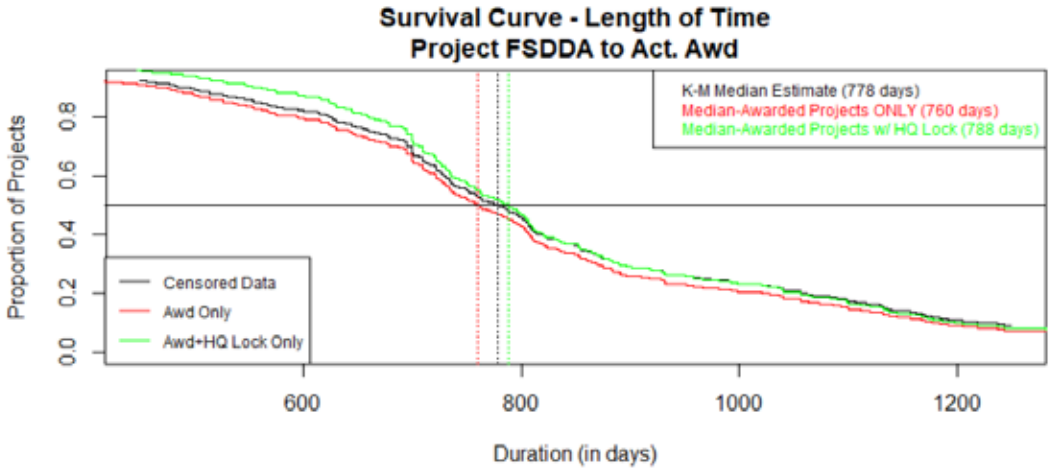

Figure 11. Survival Curve: Project FSDDA Issued to Actual Award

Table 8. Results of Survival Curve: Project FSDDA Issued to Actual Award

Number of Obs.	Number of Events	K-M Estimated Median (days)	95% Confidence		Awarded Projects Only		Awarded Projects w/ HQ Lock Date	
			LCL	UCL	Median (days)	Departure from K-M (%)	Median (days)	Departure from K-M (%)
582	443	778	752	803	760	-2.3%	788	1.3%

From Table 8, it is once again clear that removing censored data does not result in significant biasing of the design to award times. Both of the uncensored datasets result in an estimated median of design time length close to the estimate for the censored data.

3. Data Subsetting Strategy

The results of the K-M estimates provide confidence that performing an exploratory analysis on the data subsets excluding un-awarded projects is not likely to yield results that favor overly optimistic design timelines. Therefore, moving forward we perform data exploration and modeling on two subsets of the dataset independently, each corresponding to one of the two response variables of interest. Analyzing timeliness relative to the project's BY requires that projects have an Actual Award Date, but analyzing timeliness relative to HQ Execution Lock Date requires both an Actual Award Date and HQ Execution Lock Date. Table 9 illustrates the project counts by data relevant to timeliness of award.

Table 9. Available Projects per Response Variable by Project Count

Actual Award Date Listed (Y/N)	HQ Execution Lock Date (Y/N)	Number of Projects	% of Total Projects Available
Y and N	Y and N	622	100.00%
Y	Y and N	464	74.60%
Y	Y	410	65.92%

C. OVERVIEW OF ON TIME PROJECT AWARDS

For the subset of projects that are awarded prior to data collection, we mark the end of the project's BY as "09/30/ project BY." The award relative to the end of the BY is then captured as the difference between the two dates and reported as both a numeric quantity in days (– for early award, and + for late award) and a binary variable (Yes, if awarded in BY / No, if awarded after end of BY). Taken as a whole 292 out of 464 projects awarded within the individual project's BY, which is about 63%. We then apply the same method relative to the individual project's HQ Execution Lock Date, for those projects that contain both an award date and HQ Execution Lock Date. With timeliness relative to HQ Execution Lock Date, only about 45% of projects were awarded by the target date (Figure 12).

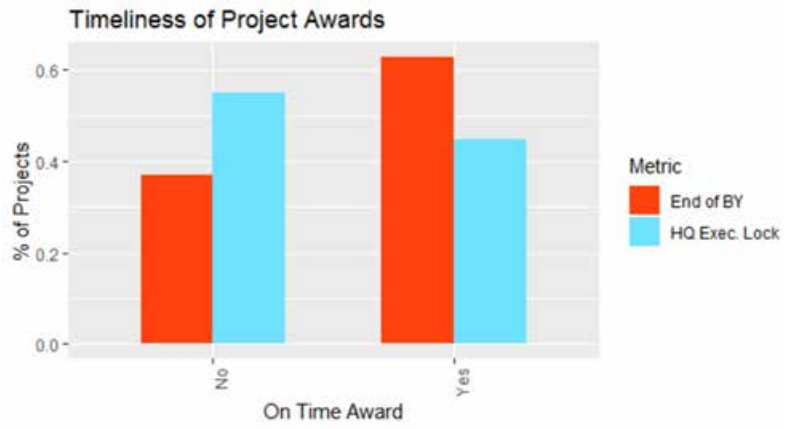

Figure 12. Proportion of On Time Project Awards

We then construct density plots using the difference between Actual Award Date and either the end of project BY (Figure 13) or the HQ Execution Lock Date (Figure 14). We add smoothers to better highlight the overall shape of the distributions. The dashed vertical line at $x = 0$, marks the end of the project year. The distribution of project award relative to the end of the BY is unimodal with a positive skew. The distribution has a central tendency about a mean of +66.5 days and a median of –10 days relative to the end of the project BY and a range of values from a minimum of –336 days to a maximum of 2155 days. The variability of observations as measured by the standard deviation is 275.7 days. There are a few projects missing the end of the BY by over 1000 days (Figure 13).

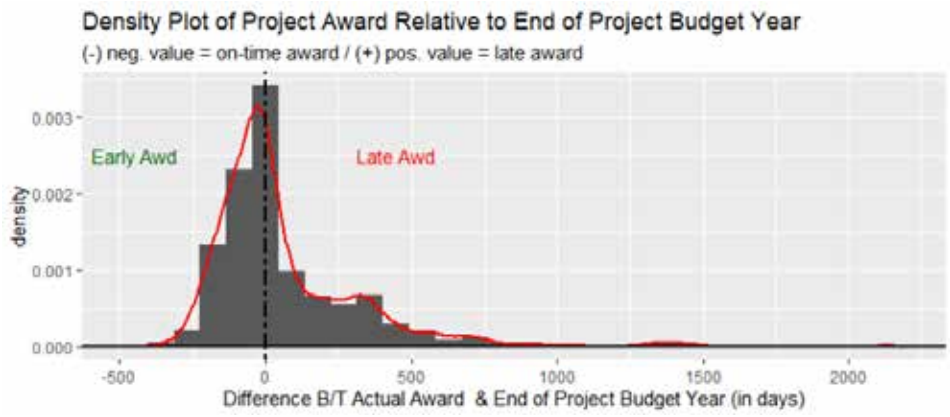

Figure 13. Award Relative to End of Project BY

Figure 14 illustrates the distribution of the project award relative to the HQ Execution Lock Date. This distribution is almost bimodal. The distribution has a mean of +37.1 days and median of +16 days relative to the individual project's assigned HQ Execution Lock Date. The range of values in this is from –281 days to 1973 days. The standard deviation is 121.3 days.

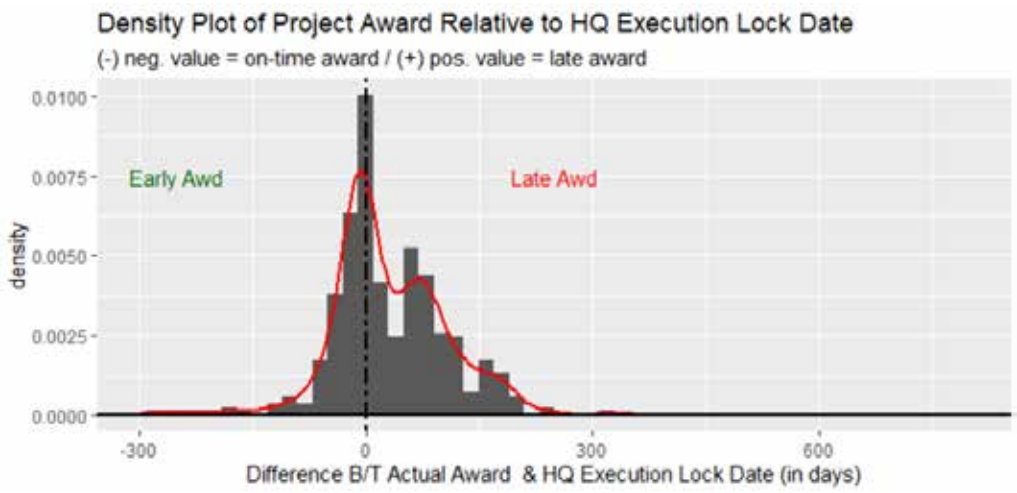

Figure 14. Award Relative to HQ Execution Lock Date

D. ANALYSIS OF COMMON PROJECT VARIABLES

In this section project performance is explored by binning projects according to three common MILCON sub-categories.

1. Design Level

One of the most basic ways of grouping MILCONs is by DBB and DB projects. One distinction between these is the amount and type of governmental involvement and oversight in the project design phase. All MILCONs fall into one of these two categories and there are differing views and opinions as to which performs better; however, DB remains NAVFAC's preferred option. This suggests that DB projects are expected to perform better. One such way of assessing performance is timeliness of award.

For the projects with Actual Award Dates, 232 projects are DB and the other 232 are DBB. For the projects featuring both an actual award and HQ Execution Lock Date, 203 projects are DBB and 207 are DB, approximately evenly split. Figure 15 illustrates on time award performance across these two project groupings.

Figure 15. Proportion of On Time Awards by Design Level

Besides measuring whether a project award occurred on time (Yes or No), we capture project performance as a duration (in days) between the actual award of the project contract and the end of the BY and/or the HQ Execution Lock Date. Ideally, that duration is negative showing that project awarded prior to either deadline. Figure 16 displays the distributions for project awards relative to the two deadlines and design level categories with summary statistics provided in Table 10.

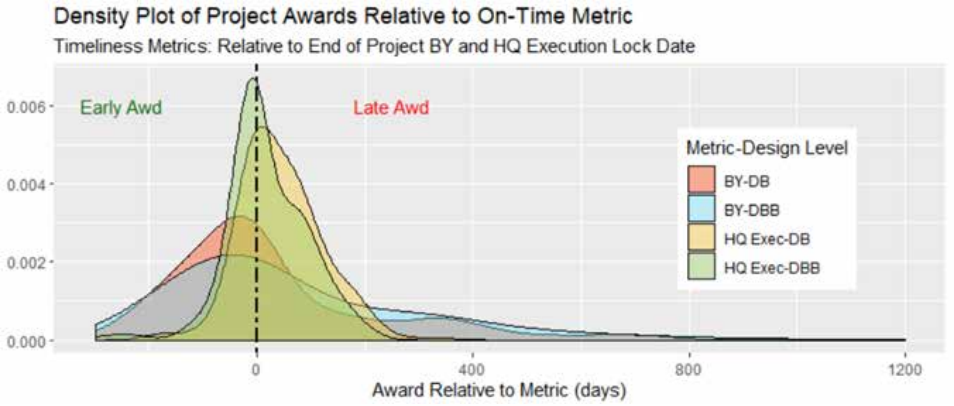

Figure 16. Timeliness of Project Award by Design Level

Table 10. Summary Statistics of Project Award Timeliness by Design Level

Summary Statistics of Project Award Timeliness (values reported in days)			
	Mean	**Median**	**Std. Deviation**
BY-DB	35.5	-13.0	223.0
BY-DBB	97.5	-4.0	317.0
HQ Exec-DB	44.6	31.0	74.3
HQ Exec-DBB	29.4	0.0	155.0

Of particular interest is the variability differences between the two performance metrics. Project performance relative to the end of the project's BY is far more variable than the distribution of performance relative to the HQ Execution Lock Date.

2. Region of Execution

The combined dataset has NAVFAC executed MILCONs from all over the world, divided into fifteen distinct regions. Each region operates with a certain autonomy but all operate under common operating procedures, guidance, and oversight. Therefore, a project should perform equally well regardless of the executing region.

Figure 17 shows individual region performance as measured against the end of a particular projects BY.

Figure 17. Navy/Marine Corps Region Performance (Project BY)

Not all regions perform equally well and workload alone does not explain away the performance differences. The three regions with the highest project workloads – NR Midlant, MCI West, and MCI East – all award projects on time over 66% of the time. Clear under-performers such as NR EURAFSWA and MCICOM/MCNCRC have moderate to low workloads but fail to award within the target BY most of the time. The box-plots in Figure 17 show the distributions of the actual project award date relative to the end of the projects BY with the dashed line marking the end of the BY. Awards to the left of the dashed line are early awards and awards to the right show a late award.

Figure 18 shows project award performance relative to the project specific HQ Execution Lock Date. One project awarded more than 1850 days after its HQ Execution Lock Date, and so, is removed to enhance plotting.

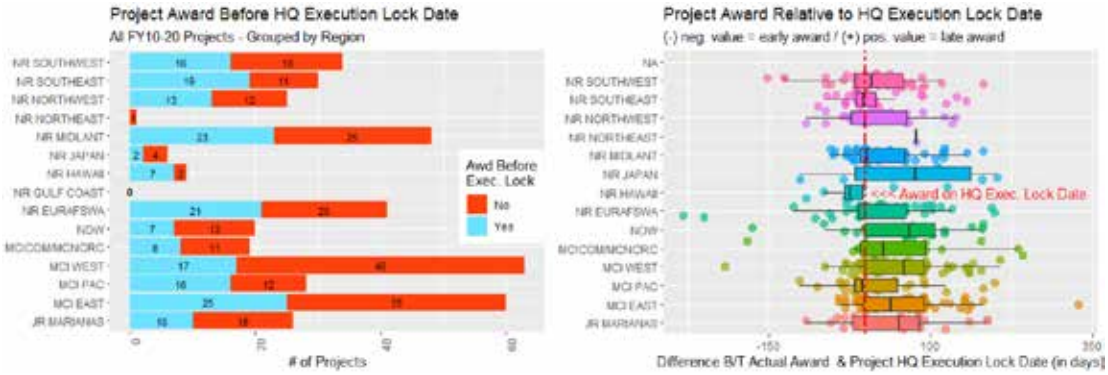

Figure 18. Navy/Marine Corps Region Performance (HQ Lock Date)

Regions perform worse across the board in meeting this delivery date. The box plots in Figure 18 shows awards relative to the HQ Execution Lock Date measured in days. The dashed line shows a contract award on the HQ Lock Date, an award left of the line shows an early award, and to the right shows late awards. As compared to awards relative to the end of the BY, there is less variability in the project awards relative to the HQ Execution Lock Date, though the central tendencies of regions show typically late awards.

3. Project Dollar Amount

MILCONs differ in cost, from a few million dollars to several hundred million. The total dollar amount will drive the level of oversight and amount of allocated resources applied towards the project's development. Therefore, binning projects by total dollar amount and then comparing relative performance across cost categories shows whether certain business practices applied at different dollar thresholds are more successful than others as measured by timeliness of contract awards.

Both EPG and eProjects report total projects costs in the individual project's BY dollars; e.g., a project with a BY of 2019 is reported in FY19 dollars. Therefore, we standardize all total project costs to FY20 dollars using the NAVFAC Building Cost Index (BCI). The NAVFAC BCI uses monthly conversion factors instead of an annual factor. We use October indices along with the following equation (Whole Building Design Guide 2020):

$$\boldsymbol{Cost}\ (\boldsymbol{FY20}\ \$) = \boldsymbol{Cost}\ (\boldsymbol{Project\ BY}\ \$)\ \mathrm{X}\ \frac{\boldsymbol{FY20\ Oct.\ Index}}{\boldsymbol{Project\ BY\ Oct.\ Index}}.$$

Once all total project costs are standardized to FY20 dollars, projects are then grouped into dollar thresholds established in coordination with NAVFAC HQ SMEs (Meek 2019a). Figure 19 shows the on time award performance per dollar threshold group and appears to indicate a possible relationship between project dollar amount and the likelihood of an on time award. As the total project cost increases, the proportion of projects that award on time relative to a project's BY tends to increase. The relationship between project cost and on time award relative to the HQ Execution Lock Date is less clear.

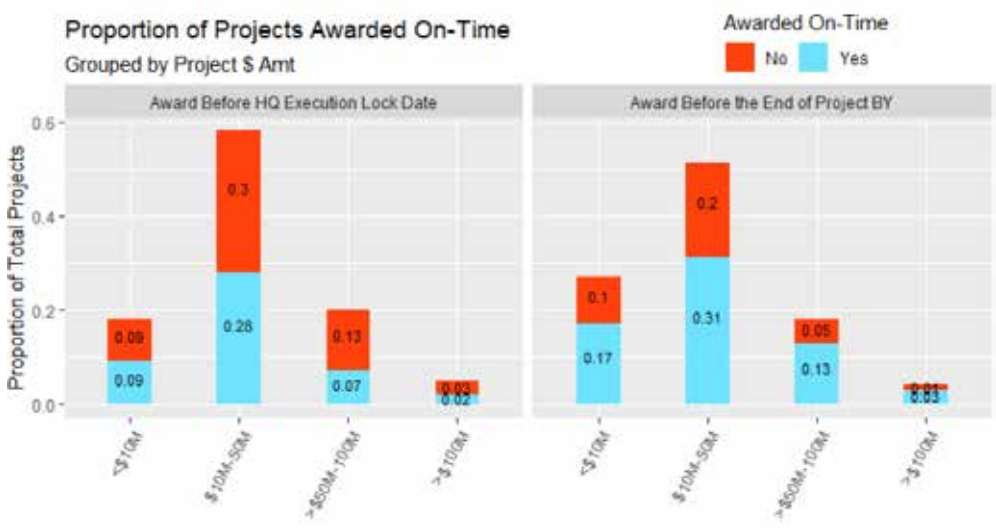

Figure 19. Proportion of On Time Awards by Project Dollar Amount

Next, we measure the difference between actual contract award date and the end of a particular project's budget (in days). The distributions for each cost group are then plotted for comparison (Figure 20). Each of the four distributions are positively skewed and feature negative median values.

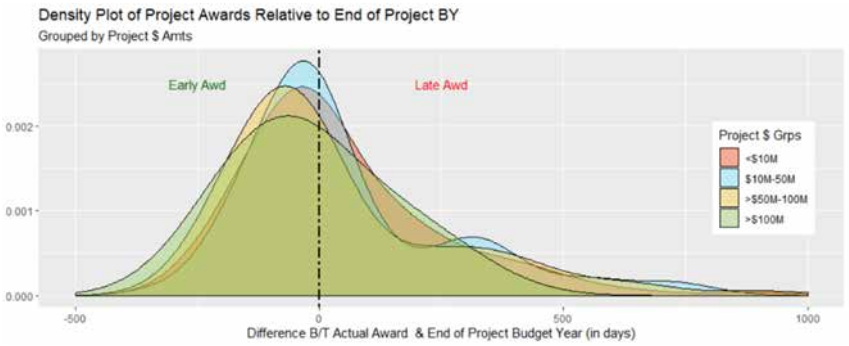

Figure 20. Distribution of Awards within Project BY by Dollar Amount

From Table 11 we can see that the central tendency of all four cost groups is to award on time, or before the end of the BY. The highest cost category features the smallest variability of any cost category.

Table 11. Project Awards Relative to End of BY by Dollar Amount

Summary Statistics of Project Award Timeliness *Relative to End of Project BY*			
Proj. $ Amts (FY20$)	**Mean** (days)	**Median** (days)	**Std. Deviation** (days)
<$10M	49.2	-6.0	217.0
$10M-50M	78.3	-7.0	278.0
>$50M-100M	76.5	-24.5	355.0
>$100M	-14.7	-8.0	159.0

The same analysis is performed for projects relative to the project HQ Execution Lock Date (Figure 21).

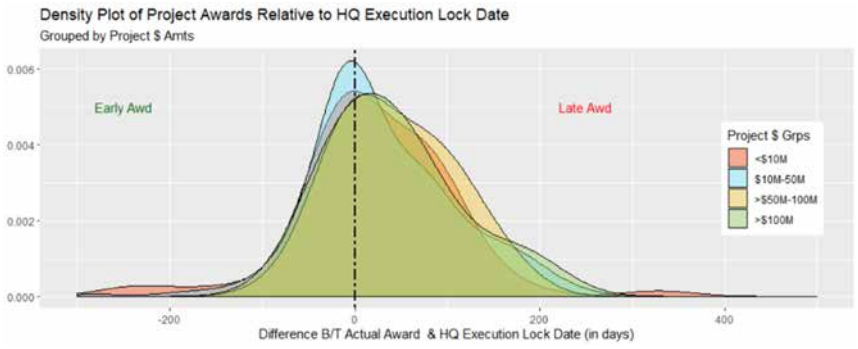

Figure 21. Distribution of Awards Before HQ Lock by Dollar Amount

The distributions associated with the four groups are again unimodal, but this time all distributions have a central tendency indicating late awards. This re-affirms previous findings that projects underperformed against this metric. The $50-100M group perform the worst with a median award time of +42 days past HQ Execution Lock Date and a substantial variation as measured by a 221-day standard deviation (Table 12).

Table 12. Project Awards Relative to HQ Lock by Dollar Amount

Summary Statistics of Project Award Timeliness *Relative to HQ Execution Lock Date*			
Proj. $ Amts (FY20$)	**Mean (days)**	**Median (days)**	**Std. Deviation (days)**
<$10M	23.3	10.0	81.5
$10M-50M	30.4	8.5	76.3
>$50M-100M	66.9	42.0	221.0
>$100M	42.8	31.5	68.7

E. ENHANCED PROJECT VARIABLES

Combining EPG and eProjects provides the opportunity to enhance the analysis of project performance by using the project data for every MILCON project dating back to 2002. Since EPG serves as the repository of all DD 1391s, various project aspects can be explored, generalized, and mapped back to any project of interest. In this section, three opportunities for dataset enrichment are explored and we assess their possible influence on the response variables of interest.

1. Primary Project Category Code Unit Cost Variability

Every MILCON project is the aggregate of several distinct construction elements and activities. Unique construction elements are identified by a category code. For example, the construction of a road has a category code of "85110." In EPG, each project appears as a collection of these category codes and the associated cost for this element. The project documentation will specify a unit of measure (UOM) and unit cost for each category code. The category code associated with the major work element, as measured by total cost or work effort, is assigned as the project's primary category code. Since late changes to project scope or cost typically result in an impact to the project's award, the relative variability of the primary category code is a logical metric to explore. If a primary category code is highly stable, then it is likely to be a well-known and often practiced work effort, resulting in a low likelihood of late project changes. Conversely, if a primary code is highly variable, this may suggest that this work effort is not well understood or planned for and so a late project change could be expected.

To explore this, we measure the variability of the unit cost across all projects, for all primary category code/UOM pairs. The 464 projects in our dataset correspond to 177 unique primary category codes, which we use to collect any record in EPG with a matching primary category code. This results in a capture of 86,593 records that contain these category codes, dating back to 2002. We then group the records by category code/UOM pairs, and compute the mean and standard deviation of the unit cost for each pair. We use the coefficient of variation (CV) for each category code and UOM pair to normalize the variation. We then map these values back on to the project details for the 464 projects by merging upon the projects primary category code and UOM. We then group the projects into "Low variability (CV<0.75)," "Medium variability ($0.75 \geq CV \leq 1.5$)," and "High variability (CV>1.5)" and plot project performance relative to our two response variables of interest (Figures 22 and 23).

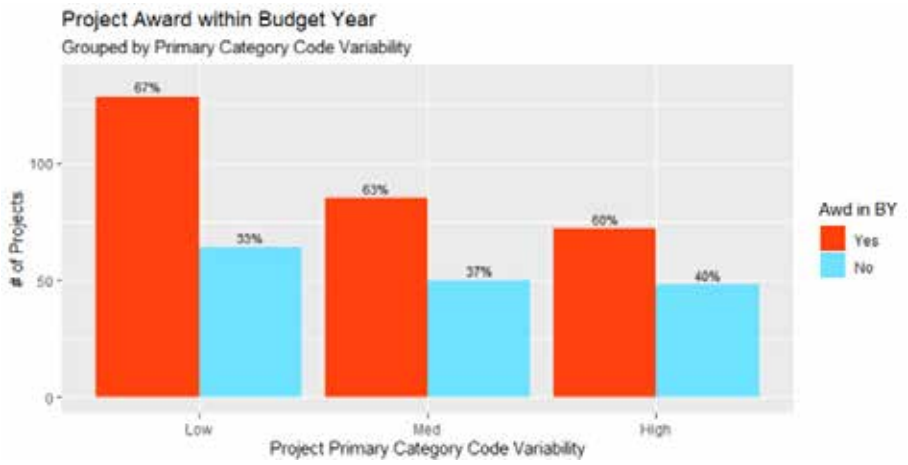

Figure 22. Award in Project BY Grouped by Category Code Variability

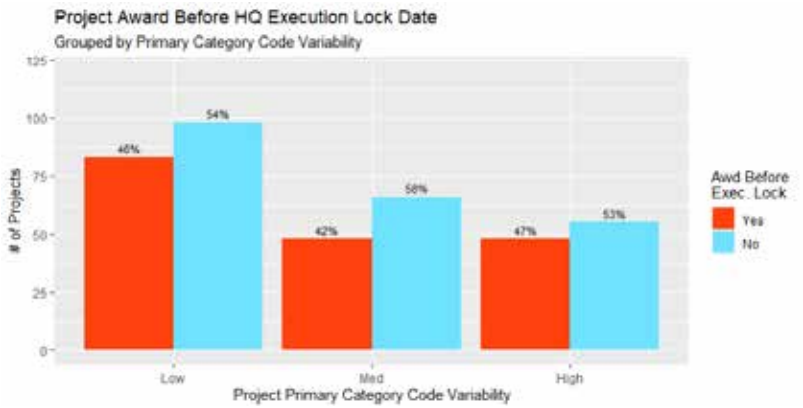

Figure 23. Award Before HQ Lock Date Grouped by Category Code Variability

The possible effect of the category code variability is clear when looking at awards within a project's BY. Those projects with a "Low" category code CV are more likely to award on time as compared to the other variability categories. The possible effect of this variable is not discernible when looking at performance relative to the HQ Execution Lock Date.

2. Variability of Total Project Cost during Development

During project development, changes to scope and cost are common. To determine the potential influence of project cost variability on the timeliness of award, the CV for the total cost of each awarded project is calculated by leveraging all "Final" versions of the project details records in EPG after normalizing all costs to FY20 dollars. Figures 24 and 25 compare the distribution of CVs for projects awarded/not awarded on time and the relative performance of projects once they are binned into Low, Medium, and High categories based on natural breaks in the CV distributions. These plots show that projects with a lower degree of variability in their total project cost during development, tend to perform better than those with higher degrees of variability.

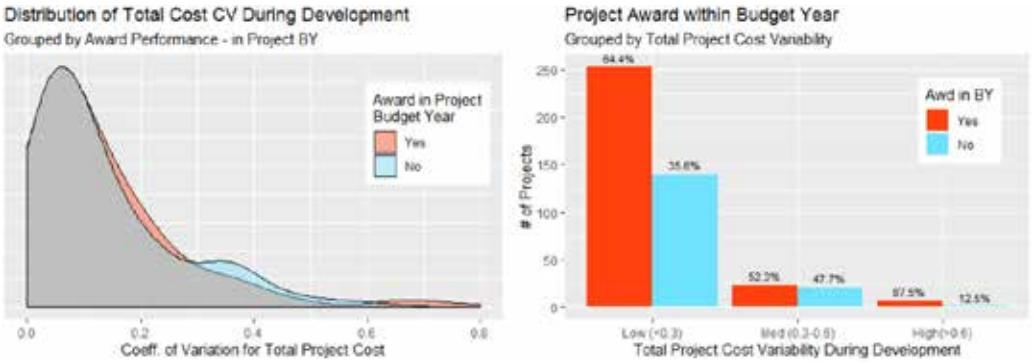

Figure 24. Awards in Project BY by Total Cost CV

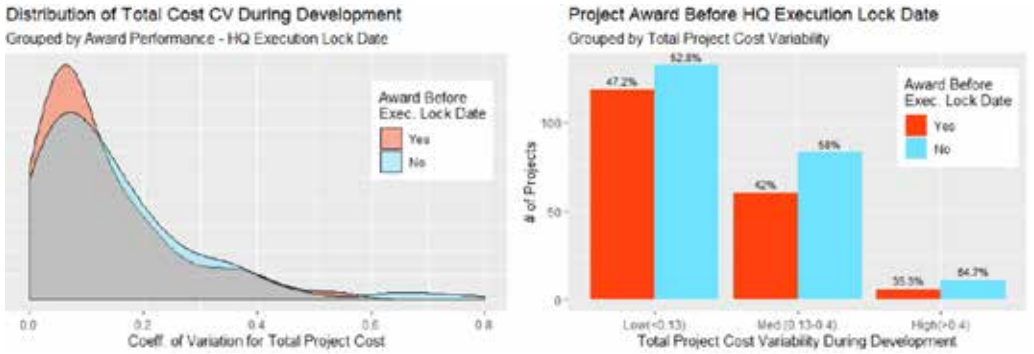

Figure 25. Awards Before HQ Execution Lock by Total Cost CV

3. Maturity of Project Documentation by Total EPG Record Count

The number and type of EPG records associated to each individual MILCON differs, despite the established workflow. Sometimes, projects receive regular updates resulting in many EPG records at various status levels. In other cases, some projects only have a handful of records. Each time a record appears, it represents a degree of work effort applied to the planning of the project. Therefore, the total number of records (at any level) for each project may illustrate project documentation maturity that should translate to fewer unforeseen changes leading up to the project award. Figures 26 and 27 illustrate award performance by number of associated project records. We group projects into Low, Medium, and High categories according to natural breaks observed in the numbers of records.

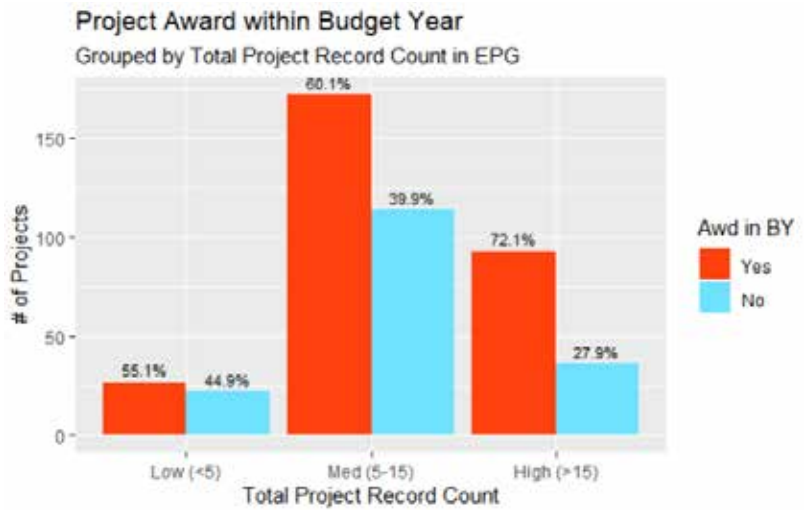

Figure 26. Awards in Project BY by EPG Record Count

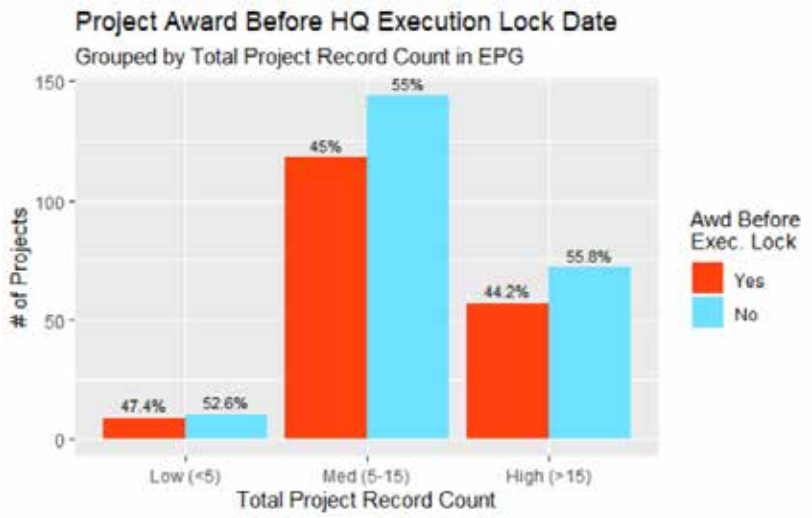

Figure 27. Awards Before HQ Execution Lock by EPG Record Count

Figures 26 and 27 show that projects with more records tend to be more likely to award on time relative to the project's BY. However, this relationship is reversed for performance relative to the HQ Execution Lock Date.

IV. MODELING AND ANALYSIS

In this chapter, we explore two machine learning algorithms to predict a project's likelihood to award on time. The intent of these models is not to predict a future project's award performance but rather to determine which project factors are most indicative of desirable project performance. For the modeling in this chapter, we only explore the categorical response variables indicating award on time or not, since we are most concerned with whether a project simply meets the desired milestone.

A. PREDICTOR VARIABLE SELECTION

1. Core Model Predictors

From the data collection and analysis of the previous chapters, we select a subset of 28 project factors from the combined EPG and eProjects dataset to serve as the core collection of predictors for modeling each of the two response variables. These factors represent either unique project characteristics or project factors that the analysis of the previous chapters found to possibly influence project performance. Project factors from EPG come from the project record with the most senior "Project_Details_Id." Table 13 summarizes the core project factors in their initial configuration prior to any changes made during the iterative model development process.

Table 13. Core Modeling Dataset

Core Modeling Dataset		
Variable	**Type**	**Description**
PROJECT_DETAILS_ID	Numeric	PROJECT_DETAILS_ID corresponding to the most senior EPG record for each particular project
Region	Factor (12 levels)	Project Region categorized as: JR Marianas, MCI East, MCI Pac, MCI West, MCICOM/MCNRC, NDW, NR EURAFSWA, NR Midlant, NR Northwest, NR Southeast, NR Southwest, Navy_Other
Execution Team	Factor (8 levels)	Project Execution Team categorized as: ML Core, NAVFAC LANT Core, NAVFAC PAC Core, NW Core, SE Core, SW Core, Wash Core, Other_ExecTeam
Design Level	Factor (2 levels)	DB or DBB
Design Agent	Factor (2 levels)	AE or IH
Design/Construction by Other Agency	Factor (2 levels)	Yes or No
Acq Tool	Factor (7 levels)	MAC-General, Stand Alone Construction Contract, MAC-8A, MAC-DB, MAC-Industrial, MAC-SB, or other
Acq Method	Factor (6 levels)	RFP-LP,RFP-BVSS (One Step), RFP-BVSS (Two Step), Sole Source, IFB/ICB, Other_AcqMethod
Std Cost FY20	Numeric	Total Project Cost Standardized to FY20 $
UOM	Factor (2 levels)	Primary Project Unit of Measure: LS or other
Cat_num	Factor (3 levels)	Number of Primary Cat Code changes during DD1391 development: 1, 2 , 3
Cat_cv	Numeric	Coefficient of variation for the unit cost of the project's primary project Cat Code/UOM pair
Record_ct	Numeric	Simple count of number of EPG records for subject project
Cost_cv	Numeric	Coefficient of variation of total project cost during project development
PDA_mth	Factor (4 levels)	Qtr of CY that PDA was given
PDA_yr_adv	Numeric	Years in advance of project BY that PDA was given
PDA_to_FSDDA	Numeric	duration of PDA to FSDDA (days)
FSDDA_mth	Factor (4 levels)	Qtr of CY that FSDDA was given
FSDDA_to_DS	Numeric	duration of FSDDA to DS (days)
DS_mth	Factor (4 levels)	Qtr of CY that DS occured
DS_to_DR	Numeric	duration of DS to DR (days)
DR_mth	Factor (4 levels)	Qtr of CY that DR issued
DR_to_RFP	Numeric	duration of DR to RFP (days)
Mission Status	Factor (2 levels)	Current or New
ES_LCA_Performed	Factor (2 levels)	Yes or no
Cert_Official_Type	Factor (3 levels)	HQ MarineCorps, Region Commander, or other
Const_Type	Factor (2 levels)	Joint Use or Unilateral Construction
ATA_mth	Factor (4 levels)	Qtr of CY that ATA was given
Standard Design	Factor (2 levels)	Yes or no

2. Addition of Supplemental Project Characteristics

In addition to the project factors presented in Table 13, EPG contains supplemental project information in the form of a series of binary questions. Because of how EPG stores its data, these details are contained in two separate tables and are presented as "Block 12 - supplemental information," indicating where within a project's DD 1391 the answers to these questions are found. The additional data is collected in the form of two CSV files supplied by the database manager (Bryce 2020). After soliciting feedback from NAVFAC SMEs, we determine that 26 of these questions are likely indicators of project performance (DeVenecia 2020). Therefore, we develop and use the schema presented in Figure 28 to map the answers to these 26 questions to their corresponding projects in order to further enhance our dataset.

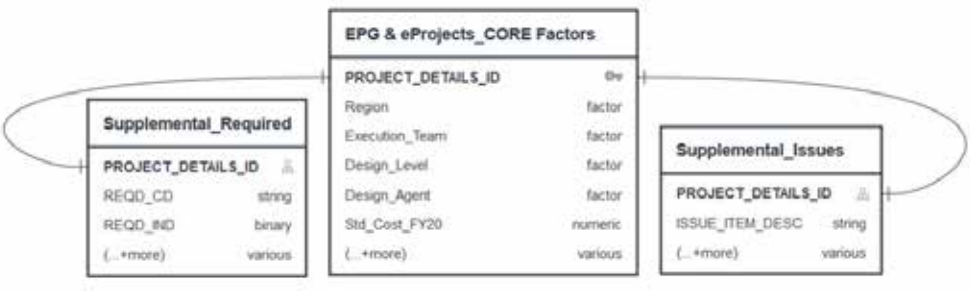

Figure 28. Schema for Inclusion of Supplemental Details

a. "Supplemental Required" table

The "Supplemental Required" table contains the same four questions for every project and an answer to these questions is required for every project. The following is a brief description of each of the four required supplemental questions with possible answer types:

1. Site approval required (Yes, Not Required, or N/A)
2. Host Nation review/approval required (Yes, Not Required, or N/A)
3. National Capital Region approval required (Yes, Not Required, or N/A)
4. Environmental cleanup required (Yes, Not Required, or N/A)

b. "Supplemental Issues" table

The "Supplemental Issues" table contains a series of specific questions in a strict binary (Yes or No) format. Within EPG, each project has a complete listing of all possible supplemental issues. The supplemental table only reports those issues that exist for the project, doing so in a transactional data format. If a particular issue appears in the dataset for a particular project, then the project is understood to have that issue. Table 14 provides a listing of the binary project issues we decide to include, grouped by overarching issue type.

Table 14. Selected Project Supplemental Details for Inclusion

Selected Project Supplemental Issues	
Issues	***Project Issues***
DDESB, AICUZ, Airfield, EMR or Wetlands	Soils-foundation and seismic conditions
Endangered species/sensitive habitat	Facility has overhead crane requirement
Known site contamination	Navy Crane Center involved
Cultural/archeological resources	Physical security
Operational problems	Local air/waster water permits required
Existing Utilities Upgrade	Historical preservation
Ordnance sweep required	Construction/operation permits required
National Environ. Policy Act (NEPA)	***Mitigation Issues***
Required documentation complete	Contaminated soil/water
Categorical exclusion	Hazardous waste
Environmental Assessment (EA)	Wetlands replacement/enhancement
Environmental Impact Statement (EIS)	
Memorandum of Negative Decision	

3. Summary of Predictors

After we map the supplemental project characteristics onto the core model predictors, we remove the "Project_Details_Id" variable from the dataset. The final collection of predictors we use for initial modeling consists of 54 project factors, 45 categorical and 9 numeric. The predictors span both EPG and eProjects and capture both project specific characteristics and various measures of performance derived in previous sections.

B. MODELING

In this section, a number of models are fit and iteratively improved in order to best determine the predictor variables that are most likely to indicate a timely project award. The goal of this section is to derive an interpretable model that can further inform ongoing NAVFAC process improvement efforts. We leverage NAVFAC SME input to identify sources of dependence amongst the selected predictors, practical importance of variables, and strategies for binning predictors into meaning full sub-categories (DeVenecia 2020).

1. Initial Modeling

Using the initial dataset established in the previous section as a starting point, simple classification trees (James et al. 2013) are fit for each of the two response variables of interest. The principal goal in doing so is to identify the relative importance of the predictors and to then identify sources of improvement to the modeling realism by further sub-categorizing, eliminating, or otherwise modify the model features.

a. Award within Project BY

Figure 29 depicts the classification tree for award performance within project BY. This tree is fit using the *rpart* package from R (Therneau et al. 2019) and by pruning to the number of splits that minimizes the cross-validated error plus one standard error.

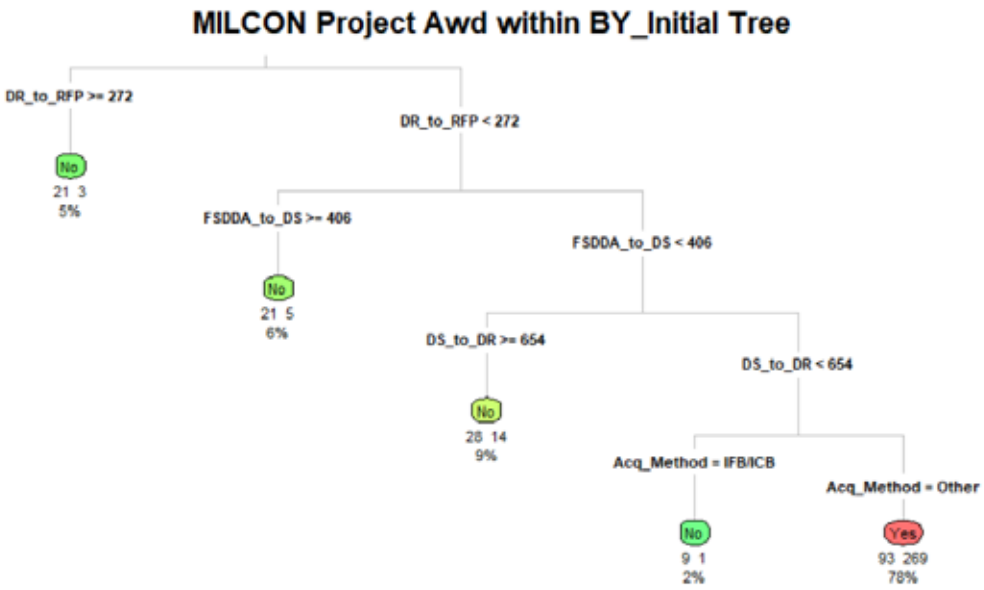

Figure 29. Initial Classification Tree for Awards in Project BY

From the model the most prominent predictor variables used to classify whether a project awarded on time relative to its BY, are those variables that measure the individual durations of the various milestones of a project's design phase. If those individual design timelines take a long time to achieve, then it is unlikely that a project awards on time. In fact, the tree in Figure 29 predicts project performance by primarily using only three design phase timelines between project FSDDA to RFP. This model successfully predicts project performance 75% of the time as evaluated against the training set. The actual splits in Figure 29 can be further used to identify prime risk indicators for project performance, though the durations used in the model for splitting are already in the extrema. Since the actual predictive capability of this model is less important than identifying possibly influential project factors for future exploration, we explore a strategy for binning the design phase timelines into more realistic sub-categories later in this chapter.

b. Award before HQ Execution Lock Date

Next, a classification tree is fit using the second response variable of interest, project award prior to the established HQ Execution Lock Date. Again the initial

classification tree is pruned in order to minimize cross-validated error plus one standard error (Figure 30).

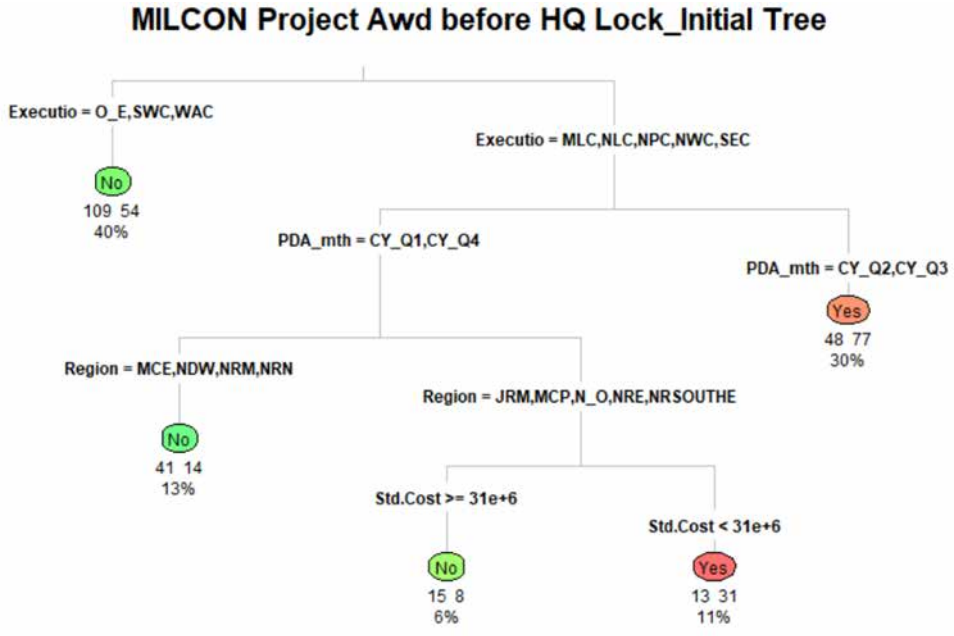

Figure 30. Initial Classification Tree for Awards Before HQ Lock Date

Unlike the results of the classification tree developed in the previous sub-section, the design phase timelines do not play a prominent role in the initial classification tree for this particular response variable. Instead both the project's execution team and region are important project factors.

2. Refinement of Predictor Variables

After completion of the initial modeling efforts, we evaluate the results of both classification trees and their corresponding lists of variable importance in coordination with NAVFAC SMEs to determine areas for improvement. We select the following two principal areas or variable groups for further refinement.

The first area identified for improvement is the weight associated to the variables corresponding to the design phase durations. Having these variables reflect the duration of their respective phases in terms of total duration (in days) results in the overshadowing of

all other variables when predicting project performance relative to the project BY. To retain the importance associated with these variables while also allowing others to contribute, we explore the variables associated with the three most prolific design phases further. The first design phase we consider is the time from Design Release to RFP. Since the project award immediately follows the RFP, the project duration measured from Design Release to RFP is too close to predicting the response; it uses a proxy variable which is essentially analogous with the response. If the duration from Design Release to RFP is very long, then it follows that the project award has a high probability of late award with little opportunity for corrective intervention by the project management team. Therefore, since the purpose of this model is to provide areas for intervention, we remove this variable from the dataset entirely. The second design phase with high variable importance in the initial models is the time from FSDDA to Design Start. Recent NAVFAC guidance establishes the project Key Performance Indicator (KPI) that if a project FSDDA to Design Start is greater than 45 days that the project is "at risk" (NAVFAC 2019a). Therefore, in the modeling dataset we change the FSDDA to Design Start variable from a numeric variable to a categorical variable with 3 levels that align with this performance trigger. Finally, we revise the variable for the project's duration from Design Start to Design Release. NAVFAC's Strategic Design 2.0 PS3.B Working Group established a goal of less than a year for this design phase (NAVFAC 2019b). Therefore, we adopt this duration as a cutoff for a 2-level categorical variable.

The second focus area we review is the relationship between a project's region and the execution team. On the surface, these variables appear to have strong dependence since the execution teams have a finite scope of responsibility that encompasses only certain regions. Therefore, the relative variable importance of one or both variables may be understated by the inclusion of both in a single model. To test this, we fit models using only one of these variables interchangeably. We find that removing one of these variables drastically decreases the model performance in both cases. This suggests that the region/execution team pairing is the important consideration. We develop a new categorical factor to capture this pairing, and remove the previous region and execution

team variables from the dataset. Table 15 summarizes the changes made to the project factors.

Table 15. Revised Predictor Variables for Modeling

Revised Predictors			
Initial Variable	**New Variable**	**Type**	**Description**
Region Execution Team	Reg_ExecTeam	Factor (16 levels)	Combined project Region and Execution Team pair
FSDDA_to_DS	FSDDA_to_DS_GRP	Factor (3 levels)	Levels: negative duration, 0-45 days, and >45 days
DS_to_DR	DS_to_DR_GRP	Factor (2 levels)	Levels: 0-365 days and >365 days
DR_to_RFP	*Removed from dataset*		

3. Model Development and Selection

In this section we fit and compare several machine learning algorithms to find a top performing predictive model for each response variable. In order to maximize utility for non-technical audiences we also consider interpretability in our final model selection. These machine learning models return the predicted probability of on time award. A project with a predicted probability above a specified threshold is predicted to be on time. For the purposes of evaluating performance across model types we compare the corresponding Receiver Operating Characteristic (ROC) curves. The ROC plots 100% minus specificity (where the specificity is the percentage of late projects correctly classified as late) against the sensitivity (the percentage of on time projects correctly classified as on time) for thresholds varying from 0–100%. Models that predict well have a ROC that increase to 100% very quickly.

In the model comparisons of this section we cross-validate the ROC curves corresponding to the random forest models. However, the ROC curves associated with the classification trees are based on the entire dataset and are not cross-validated. Instead the ROC curves are fit following the selection of the classification tree which is a result of the cross-validation internal to the *rpart* package, resulting in the selection of a classification tree that is neither under- nor over-fit. We verify this by separately conducting ten-fold cross validation of the Area Under the Curve (AUC) computed from the ROC for these

models and confirming that they are similar to the AUC's of the ROC's based on the entire dataset.

a. Award within Project BY

Using the revised dataset as defined in the previous section, we first use the *rpart* package to fit a classification tree. However, this time we prune according to two different criteria. We prune the first classification tree to the number of splits that minimizes the cross-validated error plus one standard error. We then prune the second tree to the number of splits that minimizes the cross-validated error only. Both are reasonable pruning strategies; however, pruning to minimize only the cross-validated error yields a deeper tree, providing additional insights into influential project factors worth future exploration.

Besides the classification trees, we fit two different random forests from independent R packages. We fit the first random forest using the *randomForest* package (Liaw and Wiener 2002). This package gives random forests using methods first proposed by Breiman (1996). As Breiman (1996) notes, his method tends to favor splits using variables that have a greater number of possible splits. For example, a categorical variable with ten classes has a greater chance of being chosen than one with only two classes. Thus, to facilitate interpretation we fit a second random forest using the *cforest* function from the *party* package (Hothorn et al. 2020). This function develops its random forests using conditional inference trees. This splitting criterion is not biased towards certain types of variables and can lead to model fits that are more easily interpreted.

The principal interest in exploring the random forests is to determine whether a single classification tree has sufficient accuracy compared to the random forests. If so, this would allow us to use the classification tree instead of the random forests to gather insights about the response variable. This would be preferable because of the inherent interpretability of classification trees. The ROC curves in Figure 31 compare the performance of the four models when predicting if a project successfully awarded within its BY. We note that in Figure 31 the x-axis labels are reversed to start at 100% so that they indicate specificity.

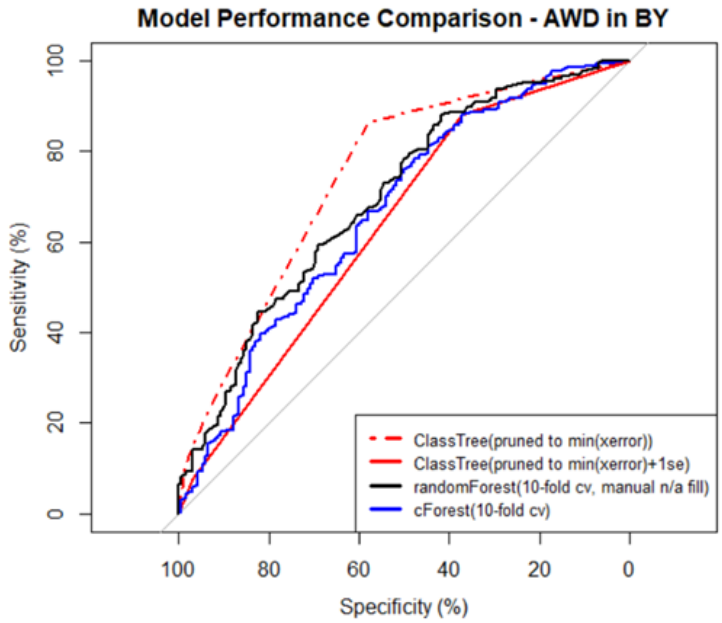

Figure 31. ROC Curve for Initial Model Comparisons (Award in BY)

As seen from Figure 31, the random forests perform about the same as one another but under-perform against the classification tree pruned to minimize the cross-validated error only. We determine this by noting the relative position of the ROC curves. Better performing models have a curve closer to the upper left corner of the ROC plot since this indicates a quicker increase to 100% sensitivity. This is an unusual outcome, since in most problems we see the random forests perform at least as well as (if not better than) either classification tree.

Random forest performance is degraded when a large number of unproductive predictor variables (i.e., noise) are included in the random forest fit. In order to then possibly improve the performance of the random forests, the dataset is further reduced to only those variables that appear as "important" in the random forest models. We present a comparison of the relative variable importance for each of the 52 predictors across the two random forest models in Figure 32.

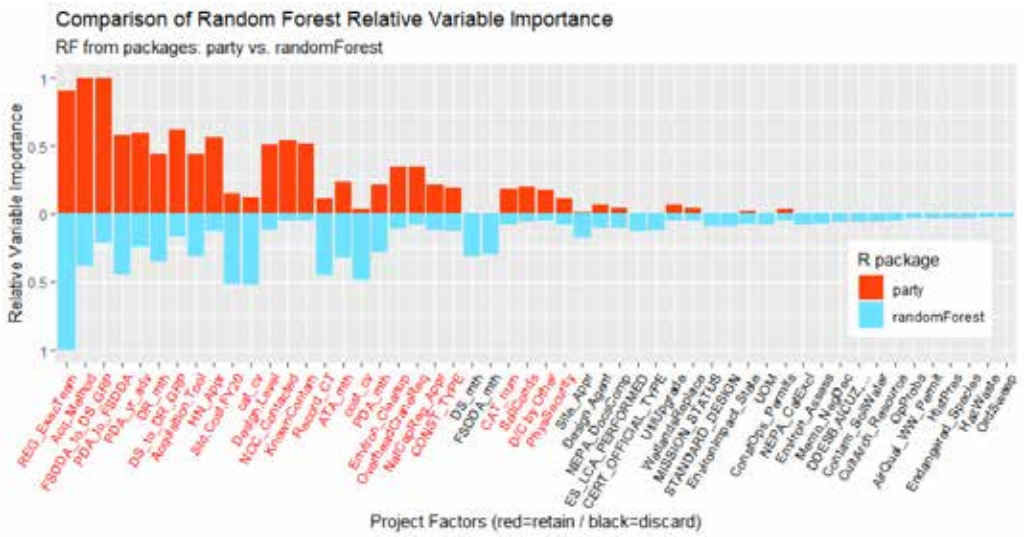

Figure 32. Comparison of Relative Variable Importance Across Random Forests (Award in Project BY)

To determine which variables are the most productive and thus should be kept for re-modeling the random forests, the variables are first sorted in order of the total relative variable importance summed across the two models. Visually, the descending rank order of variable importance can be seen by viewing the variables from left to right in (Figure 32). From Figure 32, approximately half of the predictor variables appear to be adding something to the models. Therefore, 26 of the original predictors are retained (identified in red along the *x*-axis of Figure 32). Generally, the variables kept are in the top half of the variable importance rank order, however, neither "DS_mth" nor "FSDDA_mth" are kept because of the zero importance to the random forest developed using the *party* package. Because the conditional inference splitting rule implemented in the *party* package treats variables more equitably, for our application it makes sense then to favor the variable importance from that package. Since both the "DS_mth" and "FSDDA_mth" variables have zero importance in that random forest, we skip them in favor of the next few variables.

Once we reduce the dataset to the most important variables, we again fit the random forests and compare these to the initial models. Figure 33 shows the ROC curves for all six models.

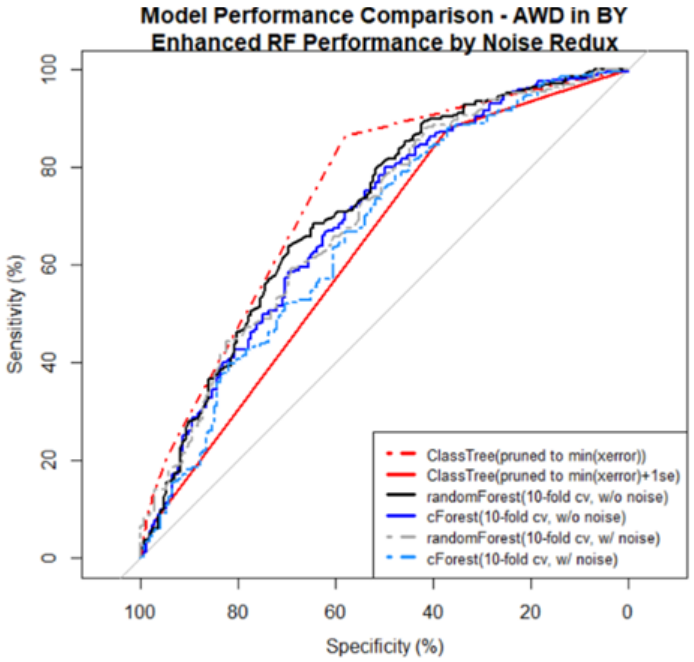

Figure 33. ROC for Final Model Comparison (Award in BY)

As anticipated, by reducing the noise in the dataset both random forests perform better overall; however, both forests are still out-performed by the classification tree pruned to minimize cross-validated error only. Because of this and because classification trees are more easily interpreted, it is then reasonable to use the best performing classification tree in order to gather insights about influential project factors that possibly drive project performance. The resulting classification tree is presented in Figure 34.

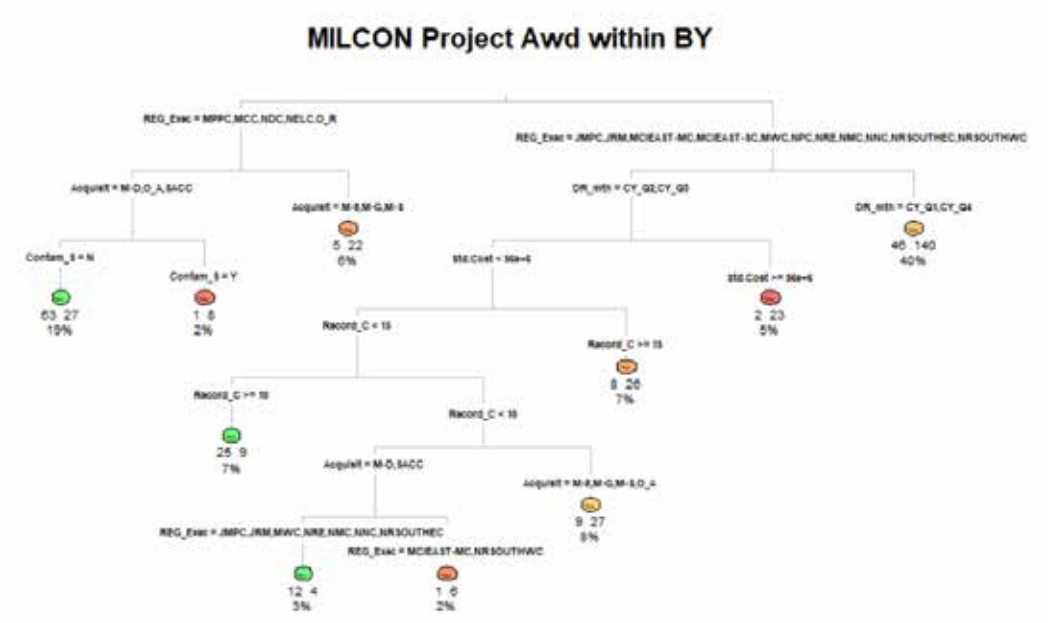

Figure 34. Selected Model for Project Awards in BY

The classification tree presented in Figure 34 has a 10-fold cross validated accuracy of 75.86%. Table 16 shows the confusion matrix using the cross-validated predictions for this tree and highlights the types of errors generated by this model.

Table 16. Confusion Matrix for Selected Model (Award in BY)

Classification Tree (Awd within BY) Confusion Matrix			
		Truth	
		No	Yes
Prediction	No	100	72
	Yes	40	252

From Table 16 there are two types of possible errors. One is when a project is predicted to finish on time but it does not, and the other is when a project is predicted to be late but instead awards on time. Of the two types of possible errors, the latter is less costly in practice. Therefore, the model is performing well in terms of minimizing the most

negative type of error since the occurrence of the most significant error type is relatively small, occurring 8.62% of the time.

The interpretability of the classification tree helps to highlight the project factors that are most significant when predicting good project performance. From Figure 34 we gather that the region/execution team and acquisition tool are some of the most important predictors of a project successfully awarding on time, in this case within its BY. The number of EPG records associated to the project, when within the calendar year the design release for a project is given, and the overall standardized cost of a project were also important factors and considerations.

b. Award prior to HQ Execution Lock Date

As with the modeling efforts in the previous section four models are fit to the refined dataset but this time with the aim of predicting performance relative to a project awarding before the prescribed HQ Execution Lock Date (Figure 35).

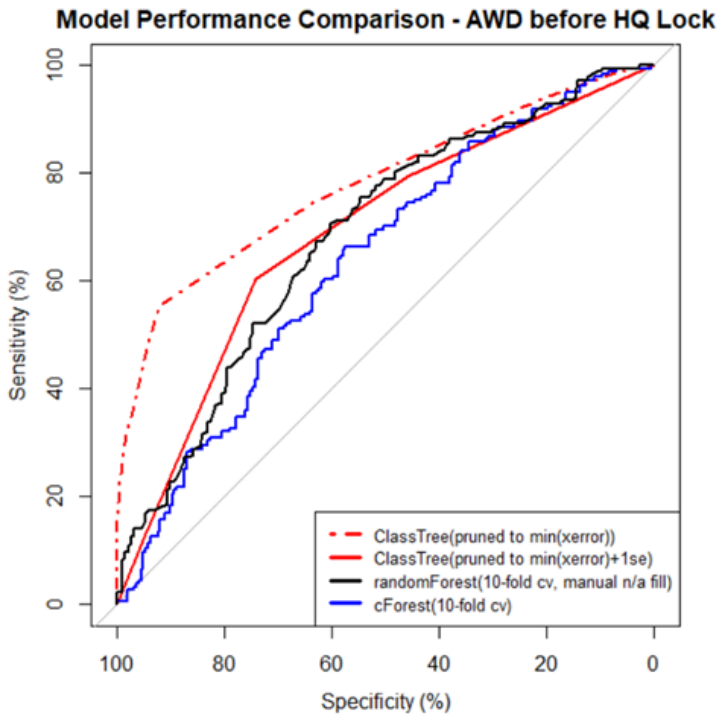

Figure 35. ROC for Initial Model Comparison (Award Before HQ Lock)

Here both of the classification trees, differing only in pruning strategy, outperform both random forest models. However, when factoring in the associated randomness of the models and the inherent instability of classification trees, the performance of the classification tree pruned to the more conservative threshold of minimizing the cross-validated error plus one standard error and the random forest generated from the *randomForest* package perform comparably well. Again, this is not what would be typically expected when comparing classification trees to random forests and so the random forest models are evaluated for unproductive variables and re-modeled to determine whether their performance could be improved. Figure 36 compares the variable importance for the two random forest models.

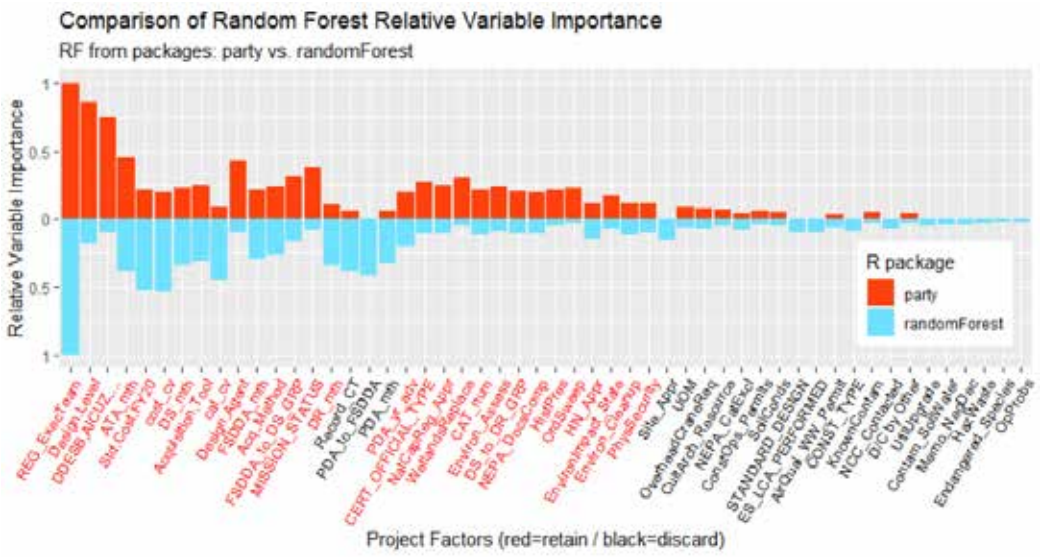

Figure 36. Comparison of Relative Variable Importance Across Random Forests (Before HQ Lock)

From Figure 36, 28 variables appear to be important across both models. Those factors are highlighted in red along the *x*-axis of Figure 36. Again owing to the differing strategies of handling categorical factors, the importance relative to the random forest generated by the *party* package is weighted heavier in determining whether or not a factor should be retained. After we eliminate the unproductive predictor variables, the random

forests are again modeled. The resulting ROC showing the relative performance of the six various models can be seen in Figure 37.

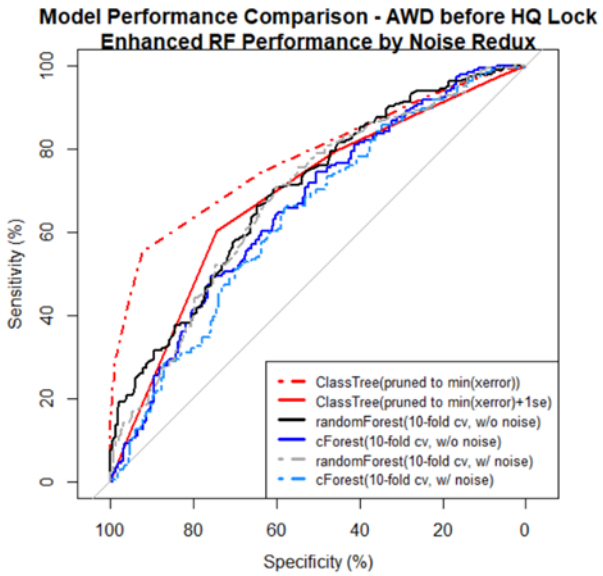

Figure 37. ROC Curve for Final Model Comparison (Before HQ Lock)

Removing the noisy variables improves the overall performance of the random forest but not enough to outperform the single classification tree pruned to minimize the cross-validated error. Therefore, using this single tree is in fact a reasonable strategy to gather insights about the response variable of interest in this section.

Figure 38 presents the classification tree pruned to minimize the cross-validated error. The resulting tree highlights the various project factors found to be the most important when predicting a projects likelihood of awarding prior to the established HQ Execution Lock Date.

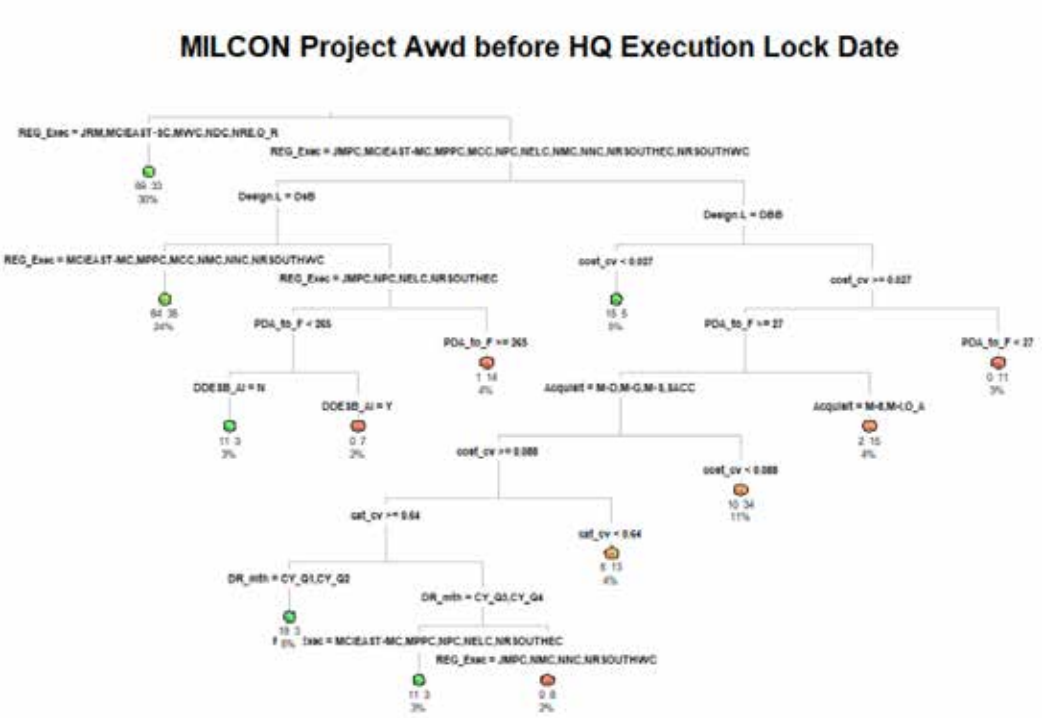

Figure 38. Selected Model for Awards Relative to HQ Lock

The selected classification tree has a 10-fold cross-validated accuracy of 75.61%. Table 17 shows the confusion matrix for this particular tree.

Table 17. Confusion Matrix for Selected Model (HQ Lock)

Classification Tree (Awd before HQ Lock) Confusion Matrix			
		Truth	
		No	Yes
Prediction	No	208	18
	Yes	82	102

V. SUMMARY AND CONCLUSIONS

This chapter presents the conclusions and findings developed over the course of our research. We start by presenting our findings pertaining to the collection and consolidation of the relevant data and identify lessons learned to enhance replicability for future research efforts. Next, we then present our findings from our initial data analysis and modeling efforts to further focus ongoing improvement efforts. The chapter then concludes with a discussion regarding future research opportunities and ideas.

A. CONCLUSIONS

1. Data Collection and Consolidation

Both EPG and eProjects are data rich resources for collecting and analyzing MILCON project performance. EPG is the key repository for project development details from initial concept up to Congressional funding, and eProjects is an important work management tool for tracking a project's performance from the point funding is received out to project award (and beyond). For future work that seeks to analyze data across the two databases, there are some issues we identify that may be of use.

At the current time, EPG and eProjects are not integrated, and so the information must first be collected and then joined across the two disparate databases. Adding complexity to this task is how EPG stores and links project details. Internal to the database, EPG uses a database specific "Project_Id" number uniquely assigned to each project. Then each subsequent record for that project receives a "Project_Details_Id," which is again unique to the internal workings of the database. EPG works as a collection of tables linked by referencing either the "Project_Id" or the "Project_Details_Id," depending on the particular tables. These labels are meaningless outside of EPG yet are the essential key for navigating, manipulating, and collecting project details. Outside of EPG, projects are most often referenced by a pairing of the projects non-unique "P" number and customer or location UIC. Fortunately, the "P" number and customer UIC are among the few data fields contained within the EPG parent table. We find that by creating a project specific "P#-

UIC" label in both the EPG parent table and eProjects dataset, we can merge project details across the databases.

Another issue is that the EPG record status levels do not conform to the more common authorization levels, i.e. Pri0, Pri1, etc. This does not affect our research but would affect a study looking at cost growth during project development, for example. As a general rule, Pri0, Pri1, and Pri2 project status levels loosely align with the following EPG record status levels (in order): INSTL/PWD_FINAL, REGN_VALID_FINAL, and CERTIFIED_FINAL.

Finally, EPG operates as a collection of project records and corresponding tables. For any project, there are often several corresponding project records with varying record status levels. For each project record, there are many corresponding tables. This means that considerable care is needed when collecting and using the data for follow-on analysis.

2. Data Analysis

Overall, our research finds that about 47% of Navy MILCON projects do not award within the planned BY and about 55% of projects do not award before the HQ Execution Lock Date. Projects tend to award within their project BY; however, this performance metric is significantly more variable than performance measured against the HQ Execution Lock Date.

Additionally, we find that a project's region, design level, and standardized total project cost all appear to have a relationship to award performance. Design Build projects tend to perform proportionally better than Design Bid Build projects in terms of awarding within a project's BY, but underperform in meeting the project's assigned HQ Execution Lock Date. Award performance relative to the project's BY is significantly more variable across both design levels than the variability in project performance relative to the HQ Execution Lock Date. Also, award performance is not uniform across regions of execution, with some regions performing relatively better across one or more of the performance metrics.

Finally, our analysis suggests that that there are other project metrics worth measuring from the EPG database in order to further enrich our understanding of project performance. For example, we find that the raw number of EPG records that a project has appears to have some relationship to how likely that project performs in terms of on-time award. We also leverage EPG to measure total project cost and unit cost variability, which also appear to add value in understanding a project's future performance.

3. Model Analysis

The models developed in our research are not good for predicting the probability that a specific project will award on time, either when measured against the project's HQ Execution Lock Date or within the designated BY. The intended purpose of our models is to identify high impact project factors that may influence the likelihood of a timely award. Though the classification trees do not establish causal relationships, they are instructive to focus future investigative analysis efforts. Table 18 shows the most influential variables for the two selected predictive models.

Table 18. Summary of High Influence Variables

Top Influential Variables Across Models		
Variable Importance Rank Order	Award in Project BY	Award Before HQ Execution Lock Date
1	**Region/Execution Team**	**Region/Execution Team**
2	**Acquisition Tool**	Variability of Total Project Cost (during development)
3	**EPG Record Count**	DoD Explosives Safety Board / Air Instal. Comp. Use Zone / Airfield / Electromagnetic Radiation / Wetlands
4	Total Project Cost (FY20 $)	**EPG Record Count**
5	Contaminated Soil or Water	Elapsed time: Preliminary Design Authority to Final Solicitation Document Design Authority (days)
6	Month Design Release Received	Variability of Category Code Unit Cost
7	Acquisition Method	Design Level
8	Known Contaminant	**Acquisition Tool**
9	Memo of Negative Decision	Authority to Advertise (month)
10	National Capital Region Approved	Preliminary Design Authority (years in advance)

From Table 18, a project's region/execution team was the most important project factor across the two selected models. Both the acquisition tool and corresponding EPG record count also appear as important in both models. Of particular interest is that five of

the supplemental project details appear in the top ten influential factors when looking at a project's likelihood of awarding within its BY. These factors primarily speak to a project's site conditions and environmental factors. Also, the measure of the relative variability of a project's total cost and category code unit cost played influential roles in whether a project awards before its HQ Execution Lock Date.

In conclusion, these factors highlight future focus areas for follow-on analysis. Our models do not say what the exact effect of each influential project factor is, only that it appeared significant when attempting to predict project award performance.

B. FUTURE RESEARCH

Our research focuses only on the analysis of project performance as it pertains to timely contract award by focusing on the pre-award timeline. However, we would recommend conducting similar research on key project performance metrics throughout the entire life cycle of a project from initial concept through project completion. The timeliness of project delivery as measured relative to the mission need by date would be a natural extension of our analysis. Beyond timelines, there are many measures of cost growth, design quality, and effectiveness of construction management that could act as response variables for similar exploration.

We would also recommend a more in-depth data analysis of EPG itself. We focus considerable effort in linking EPG to eProjects to add additional depth to our dataset, which further enhances insights into our specific problem revolving around the timeliness of project award. What our research and own exploration finds is that EPG is a rich resource of construction data spanning the last 20 years. There is tremendous potential to explore additional elements of the MILCON program and other construction types also captured by the database. For example, one initial avenue of inquiry we explore is regarding cost growth which occurs between the EPG status levels of REGN_VALID_FINAL and CERTIFIED_FINAL. As mentioned earlier in this chapter this more commonly equates to the project status levels known as Pri1 and Pri2. The EPG records corresponding with these statuses capture the current cost of a project grouped into several categories of interest to NAVFAC program managers. Therefore, the cost growth between these two authorization

levels can be measured and summarized to tease out potential areas for improvement. To prove this concept, we capture the subset of FY19 projects within EPG that have a record at each of the two record status levels. Then we calculate the cost growth between status levels, grouped by the cost category annotated within EPG (Figure 39).

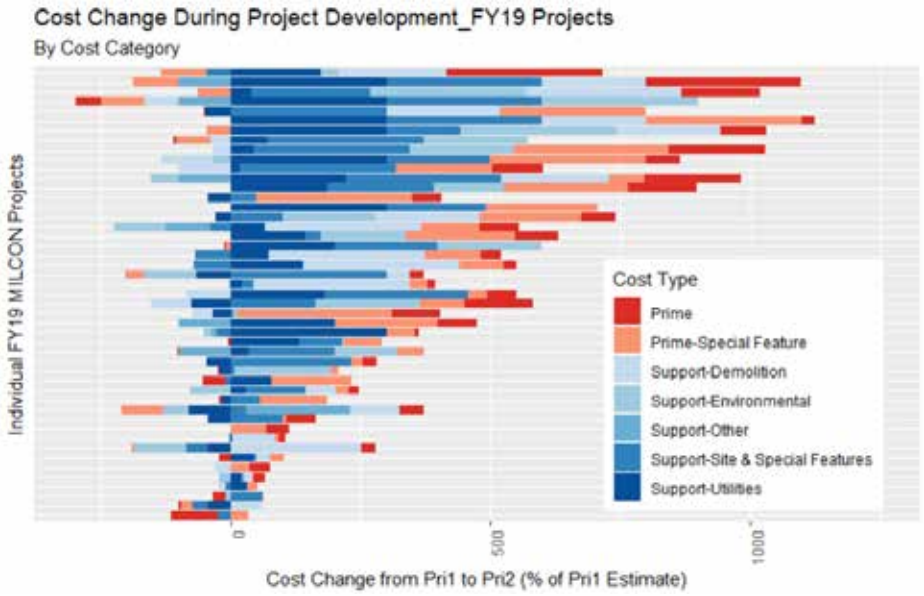

Figure 39. Example: Future Exploration of Project Cost Growth

Figure 39 shows project cost growth, grouped into principal cost categories, during the course of project development from Pri1 to Pri2 status levels. This provides program managers insight into cost growth trends across the whole MILCON program for an entire fiscal year. This in turn helps to focus process and program improvement efforts.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX. COMBINED DATASET PROJECT FACTORS

Variable	Type	Constructed	Collapsed	Number of Levels	EPG	eProjects
Budget Year	Categorical	No	No	11	Yes	Yes
Category Code - Primary	Categorical	No	Yes	Multiple	Yes	Yes
Location Unit Identification Code (UIC)	Categorical	No	No	82	Yes	Yes
UIC by Project Volume	Categorical	Yes	Yes	4	Yes	Yes
Region	Categorical	No	No	15	Yes	Yes
Responsible Component	Categorical	No	No	9	No	Yes
Responsible Team	Categorical	No	Yes	3	No	Yes
Execution Team	Categorical	No	No	23	No	Yes
Execution Team by Project Volume	Binary	Yes	Yes	2	No	Yes
Design Level	Binary	No	No	2	Yes	Yes
Design Agent	Binary	No	No	2	No	Yes
Design / Construction by Other	Binary	No	No	2	No	Yes
Acquisition Tool	Categorical	No	Yes	4	No	Yes
Acquisition Method	Categorical	No	No	8	No	Yes
Preliminary Design Authority (PDA)	Numeric	No	No	N/A	No	Yes
Final Solicitation Document Design	Numeric	No	No	N/A	No	Yes

Variable	Type	Constructed	Collapsed	Number of Levels	EPG	eProjects
Authority (FSDDA)						
PDA to FSDDA	Numeric	Yes	No	N/A	No	Yes
Design Start (DS) Actual	Numeric	No	No	N/A	No	Yes
FSDDA to DS	Numeric	Yes	No	N/A	No	Yes
Authority to Advertise (ATA)	Numeric	No	No	N/A	No	Yes
Design Release (DR) Actual	Numeric	No	No	N/A	No	Yes
DS to DR	Numeric	Yes	No	N/A	No	Yes
DR Lock	Numeric	No	No	N/A	No	Yes
Request for Proposal (RFP) Actual	Numeric	No	No	N/A	No	Yes
DR to RFP	Numeric	Yes	No	N/A	No	Yes
RFP Lock	Numeric	No	No	N/A	No	Yes
Award (AWD) Actual	Numeric	No	No	N/A	No	Yes
RFP to AWD	Numeric	Yes	No	N/A	No	Yes
Total Duration of Design Phase	Numeric	Yes	No	N/A	No	Yes
Head Quarters (HQ) Execution Lock	Numeric	No	No	N/A	No	Yes
Award in Planned Year	Binary	Yes	No	2	No	Yes
AWD to Execution Lock	Numeric	Yes	No	N/A	No	Yes
Current Status Level (CLS)	Categorical	No	Yes	10	Yes	No

Variable	Type	Constructed	Collapsed	Number of Levels	EPG	eProjects
Proposal Creation Date	Numeric	No	No	N/A	Yes	No
Number of EPG Status Levels Achieved	Categorical	Yes	Yes	9	Yes	No
Parametric Estimate Used	Binary	No	No	3	Yes	No
Energy Studies & Life Cycle Analysis Performed (ES LCA)	Binary	No	No	3	Yes	No
Originator User ID	Categorical	No	No	Multiple	Yes	No
Active	Binary	No	No	2	Yes	No
Standard Design	Categorical	No	No	3	Yes	No
Project Maturity	Numeric	Yes	No	N/A	Yes	No
EPG Routing Complete	Binary	Yes	No	2	Yes	No
Project Dollar Amount	Numeric	No	No	N/A	Yes	Yes
Standardized Project Dollar Amount	Numeric	Yes	No	N/A	Yes	Yes
Project Dollar Amount Groups	Categorical	Yes	No	4	Yes	Yes
Production Cost	Numeric	No	No	N/A	Yes	No
Other Design Cost	Numeric	No	No	N/A	Yes	No
Contract Cost	Numeric	No	No	N/A	Yes	No

Variable	Type	Constructed	Collapsed	Number of Levels	EPG	eProjects
In House Cost	Numeric	No	No	N/A	Yes	No
Total Design Cost	Numeric	Yes	No	N/A	Yes	No
Cost Growth Thru Development	Numeric	Yes	No	N/A	Yes	No
Project Element Category Code	Categorical	Yes	No	Multiple	Yes	No
Proposed Area Cost Factor (ACF)	Numeric	No	No	N/A	Yes	No
Size Adjustment Factor	Numeric	No	No	N/A	Yes	No
Original Unit Cost	Numeric	No	No	N/A	Yes	No
Unit Cost	Numeric	No	No	N/A	Yes	No
Proposed Escalation Index	Numeric	No	No	N/A	Yes	No
Guidance Escalation Index	Numeric	No	No	N/A	Yes	No
Guidance Source Description	Categorical	No	Yes	4	Yes	No

LIST OF REFERENCES

Bharat K (2019) Diagram of EPG status levels extracted from the NAVFAC MILCON primer provided to the author via personal communication, October 08.

Breiman L (1996) Bagging predictors. *Machine Learning* 24(2): 123–140.

Brown K, Dagan H, Kidda B, Ogata D (2020) Step 3: Initiate project. Lecture, MILCON Installation / PWD1391, July 14, Naval Facilities Engineering Command Process Driven Training, https://totalforcetraining.navfac.navy.mil/

Bryce C (2020) EPG datasets containing supplemental project details provided to the author via personal communication, June 12.

Bryce C (2019a) Recommendations for EPG status level consolidations and updates provided to the author via personal communication, December 02.

Bryce C (2019b) Development discussion regarding EPG data table schema conducted via personal communication, November 30.

Bryce C (2019c) EPG dataset provided to the author via email, October 09.

Department of the Navy (2018) A design for maintaining maritime superiority. Strategic Design ver. 2.0, Washington, DC. https://www.navy.mil/ah_online/ MaritimeSuperiority/

DeVenecia J (2020) Consolidated NAVFAC Headquarter SME feedback regarding supplemental project factors and suggestions for project factor handling provided to the author via personal communication, June 05.

Gondia A, Siam A, El-Dakhakhni W, Nassar A (2020) Machine learning algorithms for construction projects delay risk prediction. *J. Constr. Eng. Manage* 146(1), http://dx.doi.org/ 0.1061/(ASCE)CO.1943-7862.0001736.

Hothorn T, Hornik K, Strobl C, Zeileis A (2020) party: A computational toolbox for recursive partitioning, R package version 1.3-5. https://cran.r-project.org/web/ packages/party/party.pdf

James G, Witten D, Hastie T, Tibshirani R (2013) *An Introduction to Statistical Learning with Applications in R,* https://doi.org/10.1007/978-1-4614-7138-7.

Kaplan E, Meier P (1958) Nonparametric estimation from incomplete observations. *J. of the Am. Stat. Association* 53(282), https://doi.org/10.2307/2281868.

Larsen J, Shen G, Lindhard S, Brunoe T (2015) Factors affecting schedule delay, cost overrun, and quality level in public construction projects. *J. of Mgmt in Engineering* 32(1), http://dx.doi.org/10.1061/(ASCE)ME.1943-5479.0000391.

Liaw A, Wiener M (2002) randomForest: Breiman and Cutler's random forests for classification and regression, R package version 4.6-14. https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Mack B (2020) Improve quality, timeliness, and/or cost of DB and DBB projects. Unpublished academic paper, Clemson University, Clemson, SC.

Meek T (2019a) Guidance regarding MILCON project sub-categories and internal NAVFAC working group findings provided to the author via personal communication, February 20.

Meek T (2019b) FY10-20 eProjects MILCON dataset provided to the author via personal communication, January 16.

Naval Facilities Engineering Command (2019a) Project delivery key performance indicators. Chief's Kilo-Gram K20-02. Washington, DC.

Naval Facilities Engineering Command (2019b) Compress MILCON design process to 1 year. Strategic Design 2.0 PS3.B Working Group Charter. Washington, DC.

Naval Facilities Engineering Command (2019c) External summary memo. Strategic Design 2.0. Washington, DC. https://hub.navfac.navy.mil/webcenter/portal/Strategic_Design.

Navy Office of Information (2020) Department of the Navy FY 2021 President's Budget. Accessed July 30, 2020, https://navylive.dodlive.mil/2020/02/10/department-of-the-navy-fy-2021-presidents-budget/.

R Core Team (2019). R: A language and environment for statistical computing, version 3.6.1. https://www.R-project.org/.

Therneau T, Lumley T, Atkinson E, Crowson C (2020) survival: Survial Analysis, R package version 3.2-3. https://cran.r-project.org/web/packages/survival/index.html

Therneau T, Atkinson B, Ripley B (2019) rpart: Recursive partitioning for classification, regression and survival trees, R package version 4.1-15. https://cran.r-project.org/web/packages/rpart/rpart.pdf

Whole Building Design Guide (2020) NAVFAC Building Cost Index. Accessed July 14, 2020. https://www.wbdg.org/ffc/navy-navfac/cost-engineering-guidance-ceg/navfac-bci

Yaseen Z, Ali Z, Salih S, Al-Ansari N (2020) Prediction of risk delay in construction projects using a hybrid artificial intelligence model. Sustainability 12(4), https://doi.org/10.3390/su12041514

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California