



**NAVAL
POSTGRADUATE
SCHOOL**

MONTEREY, CALIFORNIA

THESIS

**REFINING THE PROCESS OF
THE COMMANDANT'S EDUCATION BOARD**

by

Brent J. Niewoehner

June 2023

Thesis Advisor:
Co-Advisor:
Second Reader:

Chad W. Seagren
Samuel E. Buttrey
Ruriko Yoshida

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			<i>Form Approved OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC, 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE June 2023	3. REPORT TYPE AND DATES COVERED Master's thesis		
4. TITLE AND SUBTITLE REFINING THE PROCESS OF THE COMMANDANT'S EDUCATION BOARD			5. FUNDING NUMBERS	
6. AUTHOR(S) Brent J. Niewoehner				
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) <p>As warfare becomes increasingly complex, the Marine Corps needs to better equip its warriors. One of the ways the Marine Corps does this is through graduate education programs (GEP). The Commandant's Education Board assesses candidates based on their desire, career timing, and aptitude. Aptitude is broken into competitiveness and competency. Desire is assessed via a survey. Competitiveness is assessed via fitness reports (FITREP). However, the Marine Corps has no effective method of assessing competency in the field for which a Marine is attending a GEP. This thesis used data from the Marine Corps Total Force Data Warehouse, specifically the Marine-1 and Master Brief Sheet data, to find predictors for FITREP performance of Marines filling 88xx billets. These predictors were taken from before the Marine entered into a Marine Corps GEP. This study focused on the following predictors: source of entry, time in service, primary military occupational specialty (PMOS), level of previous education (undergraduate versus graduate), and undergraduate major subject (STEM or liberal arts). This study found that source of entry, PMOS, and STEM had no predictive power in determining a Marine's performance. It also found that greater time in service and some graduate education had a slightly positive correlation to performance. This thesis recommends discontinuing vetting Marines for STEM undergraduates and giving preference to Marines with previous graduate education.</p>				
14. SUBJECT TERMS FITREPS; fitness reports; MCTFS; Marine Corps Total Force System, CEB; Commandant's Education Board, CCLEB; Commandant's Career Level Education Board, CPIB; Commandant's Professional Intermediate-level Education Board, GEP; graduate education program, STEM; science, technology, engineering, math			15. NUMBER OF PAGES 65	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

REFINING THE PROCESS OF THE COMMANDANT'S EDUCATION BOARD

Brent J. Niewoehner
Captain, United States Marine Corps
BS, United States Naval Academy, 2017

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2023

Approved by: Chad W. Seagren
Advisor

Samuel E. Buttrey
Co-Advisor

Ruriko Yoshida
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

As warfare becomes increasingly complex, the Marine Corps needs to better equip its warriors. One of the ways the Marine Corps does this is through graduate education programs (GEP). The Commandant's Education Board assesses candidates based on their desire, career timing, and aptitude. Aptitude is broken into competitiveness and competency. Desire is assessed via a survey. Competitiveness is assessed via fitness reports (FITREP). However, the Marine Corps has no effective method of assessing competency in the field for which a Marine is attending a GEP. This thesis used data from the Marine Corps Total Force Data Warehouse, specifically the Marine-1 and Master Brief Sheet data, to find predictors for FITREP performance of Marines filling 88xx billets. These predictors were taken from before the Marine entered into a Marine Corps GEP. This study focused on the following predictors: source of entry, time in service, primary military occupational specialty (PMOS), level of previous education (undergraduate versus graduate), and undergraduate major subject (STEM or liberal arts). This study found that source of entry, PMOS, and STEM had no predictive power in determining a Marine's performance. It also found that greater time in service and some graduate education had a slightly positive correlation to performance. This thesis recommends discontinuing vetting Marines for STEM undergraduates and giving preference to Marines with previous graduate education.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

1	Introduction	1
1.1	The Problem	1
1.2	Problem Significance.	1
1.3	Current Solutions	2
1.4	Proposed Solution	2
1.5	Study Contributions	3
1.6	Scope, Limitations, and Assumptions	3
2	Background and Literature Review	5
2.1	Commandant’s Education Board	5
2.2	Selection for Graduate Education	6
2.3	Graduate Education and TM2030	7
2.4	Literature Review	8
3	Methodology	13
3.1	The Data	13
3.2	Data Cleaning and Processing	13
3.3	Response Variable	15
3.4	Regression Trees	17
3.5	Random Forests	18
3.6	SVM	18
4	Analysis	21
4.1	EDA	21
4.2	Supervised Machine Learning Models	24
4.3	Linear Model	31
4.4	SVM Model of Major Subjects	34

5	Conclusions and Recommendations	37
5.1	Limitations	37
5.2	Conclusions	37
5.3	Recommendations for the CEB	38
5.4	Further Research	39
	List of References	42
	Initial Distribution List	43

List of Figures

Figure 4.1	Contains the boxplots of all the categorical variables employed in the model to make predictions. The response is the Fitness Report (FITREP) average-this-report variable created by this study. Top left is the source of entry which was either the Service Academies (ACAD), Platoon Leader’s Course (PLC), Reserve Officer Training Corps (ROTC), Officer Candidates Course (OCC), and other. Top middle is the number of years of commissioned service binned into less than 5 years, 5 to 9 years, and greater than 9 years. Top right is Primary Military Occupational Specialty (PMOS) binned into aviation, ground, or support. Bottom left is education level binned into graduate and undergraduate. Bottom right is the major subject binned into liberal arts (LA) or Science, Technology, Engineering, and Math (STEM). Data adapted from MCTFS Data Warehouse (2023).	22
Figure 4.2	The histogram plot of the response variable of FITREP average-this-report. Data adapted from MCTFS Data Warehouse (2023). . . .	23
Figure 4.3	The plots of the linear model residuals, Q-Q plot, scale-location, and the Leverage Residuals. The residuals versus fitted plot shows that the errors are roughly homoscedastic. The Q-Q plot shows that the data is roughly normal. The scale-location similarly points to the average error being zero. Finally, the leverage residuals plot shows that there are a significant subset of points that have high leverage. Data adapted from MCTFS Data Warehouse (2023).	24
Figure 4.4	The <i>rpart</i> decision tree for the average-this-report response variable broken down into five groups with the predicted response value and the percent of the population in that group. If a decision is true then data goes left. The shading qualitatively reflects the average in each node.	26
Figure 4.5	The comparison plot of the average-this-report response variable between where the <i>rpart</i> tree predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.	27

Figure 4.6	The <i>rpart</i> regression tree for the relative-value-at-processing response variable broken down into five groups with the predicted response value and the percent of the population in that group. If a decision is true then data goes left. The shading qualitatively reflects the average in each node.	28
Figure 4.7	The <i>rpart</i> regression tree for the relative-cumulative-value response variable broken down into five groups with the predicted response value and the percent of the population in that group. If a decision is true then data goes left. The shading qualitatively reflects the average in each node.	29
Figure 4.8	The comparison plot of the average-this-report response variable between where the randomForest algorithm predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.	30
Figure 4.9	The comparison plot of the average-this-report response variable between where the bagging algorithm predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.	31
Figure 4.10	The comparison plot of the average-this-report response variable between where the <i>cforest</i> algorithm predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.	32
Figure 4.11	The comparison plot of the average-this-report response variable between where the linear model's predicted response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.	33
Figure 4.12	The plot of the predicted probabilities in color scale with the source of entry on the x-axis and the response variable on the y-axis. The response here is the average-this-report FITREP variable created by this study.	34

List of Tables

Table 4.1	The table of the Root Mean Squared Error (RMSE) values of the linear model and machine learning models for the three different response variables. These values are measured from the FITREP response variable which have a minimum of 1, a maximum of 7, and an average around 4.4.	25
Table 4.2	The table of the coefficient values of the linear model created using <i>stepAIC</i>	32

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

1stLt	1st Lieutenant
ADP	Advance Degree Program
AMOS	Additional Military Occupation Specialty
BMOS	Billet Military Occupational Specialty
Capt	Captain, USMC
CCLEB	Commandant's Career Level Education Board
CEB	Commandant's Education Board
CPIB	Commandant's Professional Intermediate-Level Education Board
CSC	Command and Staff College
DOD	Department of Defense
EDA	Exploratory Data Analysis
EDIPI	Electronic Data Interchange Personal Identifier
ELP-L	Excess Leave Program - Law
EWS	Expeditionary Warfare School
FITREP	Fitness Report
FLEP	Funded Law Education Program
GEP	Graduate Education Program
GPA	Grade Point Average
JAG	Judge Advocate General

LSAT	Law School Admission Test
MBS	Master Brief Sheet
MCTFS	Marine Corps Total Force System
MOE	Measures of Effectiveness
MOS	Military Occupational Specialty
NPS	Naval Postgraduate School
OAD	Operations Analysis Division
OCC	Officer Candidates Course
PHDP-T	Doctor of Philosophy Program - Technical
PLC	Platoon Leader's Course
PME	Professional Military Education
PMOS	Primary Military Occupational Specialty
RMSE	Root Mean Squared Error
RO	Reviewing Officer
ROTC	Reserve Officer Training Corps
RS	Reviewing Senior
SEP	Special Education Program
SOE	Source of Entry
STEM	Science, Technology, Engineering, and Math
SVM	Support Vector Machine
TBS	The Basic School
TFDW	Total Force Data Warehouse

TIS	Time in Service
TM2030	Talent Management 2030
USN	U.S. Navy

THIS PAGE INTENTIONALLY LEFT BLANK

Executive Summary

As warfare becomes increasingly complex, the Marine Corps is seeking to better equip its warfighters. One of the critical ways the Marine Corps does this is through graduate education programs. Candidates for graduate education are selected by the Commandant's Education Board (CEB). The CEB is broken into two sub-boards: the Commandant's Career Level Education Board (CCLEB) and the Commandant's Professional Intermediate-Level Education Board. For both boards, the CEB assesses candidates based on their desire, career timing, and aptitude. Desire is assessed via a survey. Aptitude is further broken into competitiveness and competency. Competitiveness is assessed via fitness reports (FITREP). However, as it stands the Marine Corps has no effective method of assessing competency in the particular field for which a Marine is attending a graduate program.

This thesis used data from the Marine Corps Total Force System (MCTFS) data warehouse, specifically the Marine-1 and Master Brief Sheet data, to find predictors for FITREP performance of Marines filling 88xx billets. These predictors were taken from before the Marine graduated from a Marine Corps Graduate Education Program. This study focused on the following predictors: source of entry, time in service, primary military occupational specialty (PMOS), education level (undergraduate versus some graduate education), and finally their undergraduate major subject (STEM or liberal arts). We investigated the FITREP scores: average-this-report, relative-value-at-processing, and relative-cumulative-value. However, we based our conclusions and primary models on our response variable that we derived from average-this-report FITREP scores. This was due to a larger percentage of the population having average-this-report values than the other two, and to it having a lower root mean square error (RMSE) for its models. Our response variable was specifically a time-averaged FITREP score across a Marine's FITREPs as an 88xx.

For our models, we built a linear regression model and several machine learning models to predict FITREP performance. For the machine learning models, we built a regression tree and several ensemble models. Due to a lack of variability in our data across categories, none of our models were able to effectively predict the response variable. Some of this is due to an absence of a numeric variable, but some of it also may be due to our categories having little to no influence on our response. Finally, we built a support vector machine (SVM)

model to further study the relationship between FITREP performance and undergraduate major subjects. We wanted to more thoroughly see if STEM degrees correlated to higher FITREP scores. Our model reinforced our finding that STEM did not correlate positively or negatively in a meaningful way with our response variable. This would indicate that type of undergraduate degree is not influential in a Marine's performance as an 88xx.

Therefore, this study found that source of entry, PMOS, and major subject had no predictive power in determining a Marine's performance. It also found that greater time in service and some amount of graduate education had some positive correlation to performance, but the correlation was slight. Therefore, this thesis recommends to the board that it discontinue vetting Marines for STEM undergraduates or some level of math experience. It also recommends that preference be given to those with some graduate education.

Acknowledgments

First I would like to thank my professors Seagren, Buttrey, and Yoshida for guiding me through the thesis process, and helping over the finish line to graduation. “An intelligent heart acquires knowledge, and the ear of the wise seeks knowledge.” (Proverbs 18:15)

Doctor Major Ezra Akin (or is it Major Doctor): I appreciate your counsel and advice throughout the program and moving into the larger OR community within the Marine Corps. The time spent with you, Abby, and the kiddos last December is precious to me. “Truly, truly I say to you, whoever receives the one I send receives Me and whoever receives Me, receives the One who sent Me.” (John 13:20)

I would like to thank both Professors Robert Niewoehner for assisting me by providing me with wisdom and encouragement throughout my entire life and graduate school. Dad, I really appreciate all the ways you have loved and counseled me through my career as a Marine Corps Officer. RJ I am so thankful for all your help with Overleaf, OR, and executing my vision for awesome backcountry trips. The scriptures speak truly when it says, “Whoever walks with the wise becomes wise.” (Proverbs 13:20a)

Mom, I am so thankful for all your hard work and dedication in homeschooling me. Despite how much people tease me for being homeschooled, I finished school before noon most of my life and I still graduated from grad school! All this is because of you, Mom! You were faithful to “Train up a child in the way he should go; even when he is old he will not depart from it.” (Proverbs 22:6)

My beloved wife, Hannah, and my girls, Ellie and Lucy, I am so thankful for how you have loved me, walked with me, and done God’s Will with me. It is a joy to be a part of such a family. You all are more precious to me than I can say. As our Lord said, “A new commandment I give to you that you love one another. As I have loved you, you also are to love one another. By this all people will know that you are My disciples, that you have love for one another.” (John 13:34-35) You have indeed loved me in a Christ-like manner and for that I am profoundly grateful. Ellie, I am particularly appreciative of your desire to snuggle and wrestle that so thoroughly distracted my mind from writing this thesis when I wanted

distraction most!

Finally, I am thankful to my dearest Lord Jesus who has saved me from my sin and walked with me all my days. I believe You when You say, “I am the good Shepherd. I know My own and My own know Me. Just as the Father knows Me and I know the Father and lay down My life for the sheep.” (John 10:14-15) Lord Jesus I pray that Your glory would be made known through this work, that all may know this of Your Father, “And this is eternal life, that they know You, the only true God, and Jesus Christ whom You have sent.” (John 17:3)

CHAPTER 1: Introduction

1.1 The Problem

Entrance into Marine Corp's graduate education programs is currently administered by the Commandant's Education Board (CEB), which contains two subordinate boards, the Commandant's Career Level Education Board (CCLEB) and the Commandant's Professional Intermediate-Level Education Board (CPIB). These two boards screen, process, and select Marine Corps officers to attend various graduate programs. Each of these graduate programs are tied to specified billets that must be filled by graduates. As it stands Marines are selected for graduate education programs based on three factors: desirability, career timing, and aptitude (Foley 2022). Desirability is how much the Marine wants to attend a given program. This is measured by the CEB survey, which is sent out to all Marines who are being screened each year. Aptitude is broken into two categories: competitiveness and competency. Competitiveness is the measurement of the Marine's performance against his or her peers. Competitiveness is determined from the Marine's Fitness Report (FITREP)s and the Marine's relative performance compared to his or her peers. Competency is the measurement of a Marine's ability to successfully fill an associated utilization tour billet. Competency is the factor that is currently difficult to assess, because there is no survey, assessment, or product that can evaluate an officer for his or her competency in a given billet. Thus, there is a gap in the Marine Corps's ability to assess and select Marines for graduate education.

1.2 Problem Significance

As previously stated, these graduate education positions are tied to particular utilization tour billets. These billets are often highly technical, and cannot be filled by a regular staff officer. These billets often have high barriers to entry in mathematics, physics, engineering, computers, language, or law. Therefore, the Marine Corps must properly select Marines to attend the appropriate education program. The Marine must succeed in earning the associated degree, and then he or she will be assigned to a graduate billeted utilization tour.

These steps must be achieved to ensure the Marine Corps remains at the cutting edge of technical innovation in 21st-century warfare.

As warfare's complexity increases, the need for these highly educated officers increases. The Marine Corps has limited ability to fill technical billets with civilian experts. The issue becomes a matter of experience and expertise. However capable, these civilian experts generally do not have the experience to solve operational problems that the Marine Corps regularly faces. Thus, their ranks must be bolstered with active-duty technical experts who primarily come from a graduate education program in their technical fields. Therefore, the CEB has implications across the Marine Corps in manpower, operations, technology, acquisitions, and many more essential functions. As can be clearly seen, the competency of a selected Marine in his or her designated program is essential for the board to continue to perform its function as it selects over a hundred officers for these billets every year.

1.3 Current Solutions

As it currently stands, the CEB reconciles a Marine's desire and their competitiveness against their peers to select candidates for graduate education. However, they do not have a solution to screen Marines for their competency. For instance, Marines will be flagged as potential candidates for a technical degree if they have some level of college calculus, physics, or engineering. Therefore, a Marine with one semester of college calculus could be selected to attend a graduate program with advanced classes in differential equations, stochastics, computer science, or thermodynamics. As it can be imagined, for many programs one semester of calculus hardly recommends an individual to a master's degree in Science, Technology, Engineering, and Math (STEM). However, the reverse could be true, and previous STEM experience may be unnecessary in the creation of master's graduates in STEM. The current means of assessing graduate education competency is insufficient to meet the demands of the Marine Corps.

1.4 Proposed Solution

We apply linear regression and supervised machine learning techniques to find predictors that indicate success in a graduate education program associated with a utilization tour. We use FITREP scores as our response variable. Our predictors come from the Marine Corps

Total Force System (MCTFS) (MCTFS Data Warehouse 2023). For this research, we are particularly interested in undergraduate education subjects, Primary Military Occupational Specialty (PMOS), education level (undergraduate only or some graduate), source of entry, and time in service to see their influence on performance as an 88xx.

1.5 Study Contributions

We aim to contribute questions that can be incorporated into the CEB survey. These questions will enable the board to better gauge which Marines are suitable candidates for graduate education. These questions are informed by the traits or groupings of traits mentioned above. They are the board's means of discovering the traits that could indicate technical competency.

Additionally, this study contributes to the Marine Corps's overall understanding of the candidates that it sends to graduate education. This understanding can help improve billet assignments and talent management. Both of which contribute to the Marine Corps's appeal as a worthwhile career to talented professionals.

1.6 Scope, Limitations, and Assumptions

This research focuses on graduate education that supplies the Marine Corps with its 88xx Additional Military Occupation Specialty (AMOS) Marine Officers. The data starts with Marines who earned an 88xx AMOS in 2007 and goes through 2021. We are not adjusting average-this-report FITREP scores for time, but rather assuming Reviewing Senior (RS)s assign FITREP scores equitably over the course of their careers. We assume the data is accurate and free of error.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 2: Background and Literature Review

In this chapter, we describe the CEB, walk through the selection of Marines to graduate education, and discuss pertinent literature to this study. The literature review includes Talent Management 2030 (TM2030), other graduate theses relating to graduate education and the military, and similar studies conducted by civilian institutions to find key predictors to success in graduate education.

2.1 Commandant's Education Board

The CEB is comprised of two boards; the CCLEB and the CPIB. The CCLEB selects officers, primarily 1st Lieutenants and occasionally junior Captains, for graduate education. Its sister board the CPIB selects officers, primarily senior Captains and sometimes junior Majors, for graduate education. These two boards select officers for a variety of graduate education programs. The most notable are the Professional Military Education (PME) graduate education programs: Expeditionary Warfare School (EWS) and Command and Staff College (CSC). These two programs receive a heavy emphasis from the Marine Corps due to their direct applicability to the warfighting profession. Other graduate education programs include the Special Education Program (SEP) and the Advance Degree Program (ADP). The SEP is the Marine Corps program to send officers to receive technical degrees from the Naval Postgraduate School (NPS). ADP is its sister program where officers go and receive technical master's degrees from civilian universities. The Funded Law Education Program (FLEP), Excess Leave Program - Law (ELP-L), and the Doctor of Philosophy Program - Technical (PHDP-T) are separate but related boards for sending Marines to graduate education. The FLEP and the ELP-L are the Marine Corps two means of commissioning lawyers into the active duty service as Judge Advocate General (JAG)s. Therefore, most higher education opportunities pass through the CEB making it an incredibly important entity for future Marine Corps success. This board meets once a year to make these selections.

2.2 Selection for Graduate Education

In order to ensure success in these new technical fields the Marine Corps has established boards to assess and select officers for them. These boards aim to choose officers with the highest potential for these graduate education programs. They do this by assessing officers for the following qualities: desire, career timing, and aptitude (Foley 2022). The desire is the stated desire of the assessed Marine. The board will prioritize the Marine Corps over the desires of the Marine. However, to as great an extent as possible, the board seeks to do what is best for the Marine. What is best for the Marine is often voiced in their desires, so this is considered whenever possible when vetting candidates. Aptitude is broken down into two attributes; competitiveness and competency. Competitiveness means competitiveness against one's peers. This is assessed via a Marine's FITREPs. In a FITREP a Marine is compared against his peers in the profile of the RS and the Reviewing Officer (RO). Thus, the nature of the reviewing system is to evaluate for competitiveness against one's peers. The final attribute of competency refers to the Marine's ability in the program to which the board is considering assigning them. For instance, the board endeavors to assign Marines who have some STEM experience to STEM graduate degrees. As it stands, the board has no satisfactory or comprehensive means by which to assess Marines for this final attribute.

Before being selected by the board, Marines packages are assembled by the Manpower and Reserve Affairs graduate education monitor. The monitor screens the desirability survey, the Marine's career timing, their FITREPs, and finally their academic history to build a package for each Marine. Depending on specific traits in these packages, Marines are recommended for various graduate education programs. For instance, a Marine who does not want to attend graduate education is generally not recommended. Another example is a Marine with previous STEM experience might come highly recommended to the board to attend a SEP. Finally, a Marine with a high score on the Law School Admission Test (LSAT) and acceptance to a law school would be recommended as a candidate for FLEP. All of these packages are built by the above graduate education monitor for ease of selection by the board.

This is the feature space to which this study contributes. However, our goal is to take it a step further. The Marine Corps not only needs Marines who will succeed in their assigned program of study, but also they need these Marines to be successful in the billets

associated with their degrees that they fill after earning their degree. For instance, it is of minor importance that a Marine Operations Research student earn a high Grade Point Average (GPA) at NPS, but it is of great importance that that Marine be an excellent analyst upon being assigned to Operations Analysis Division (OAD) post-graduation.

2.3 Graduate Education and TM2030

As war becomes increasingly complex the tools with which war is waged become more and more complex. These new warfighting domains like Information Operations and Cyber Operations require levels of education unnecessary in previous conflicts. Computers are now ubiquitous across the service and their applications are increasingly diverse. This requires that the Marine Corps have technical experts and warfighting professionals in a variety of fields to achieve the tactical and operational advantage over its adversaries. It is no longer sufficient to have the one Lance Corporal in a Company who is “good with computers” to wage modern warfare. Logistic node problems and ship ballistic missile defense problems cannot be solved by hand by a planner who has a good sense for logistics or defense in depth. These are advanced problems that require advanced degrees to understand the mathematics and science behind their implementation. Graduate education seeks to solve these problems by providing the Marine Corps with technically competent warfighters and researchers.

In 2021, Commandant of the Marine Corps General Berger released TM2030. General Berger stated, “Our organization, processes, and approach to personnel and talent management are no longer suited to today’s needs and incompatible with the objectives of Force Design 2030...This report charts a new course for our personnel system.” (Berger 2021) This has significant implications for graduate education, due to it being the training ground for many of the Marine Corps technical staff billets.

On the most basic level, graduate education accomplishes one of the goals of TM2030; it matures the force. By enticing Marine Officers with graduate degrees, who would otherwise get out of the Marine Corps, it ages the force by keeping key talent within its ranks. Additionally, due to the career timing of the CPIB, it often captures senior Captains and Majors for a career of twenty or more years.

Graduate education also achieves the stated goal from TM2030 of “creating a path for

talented staff officers.” (Berger 2021) It also explicitly states, “Making all officers conform to fit industrial age notions of leadership inefficiently matches the talent of our officers with appropriate billets, impacting the effectiveness of our commanders and their staffs. It also results in the early separation of talented officers who seek staff leadership roles but see no opportunities for upward mobility.” (Berger 2021) Graduate education utilization tours can and often do lead to subsequent utilization tours in specialized fields. This allows the Marine Corps to keep officers who are no longer interested in serving in their PMOS. Thus, a talented infantry officer, who is no longer interested in filling operational billets in the infantry, can still achieve career success through graduate education and its associated billets.

Additionally, graduate education allows for the Marine Corps to achieve the goal from TM2030 of “prioritizing data and employing better analytics.” (Berger 2021) The active duty military analysts doing this stated analysis often come from Marine Corps graduate programs. Therefore, selecting high-quality officers who want to do military analysis is essential for achieving this TM2030 goal. It also states that it wants “decision support tools for promotion and selection boards.” (Berger 2021) These decision support tools are made by military officers who are currently or were previously students of Marine Corps graduate education programs. Creating decision support tools is clearly a role for these officers.

Overall, TM2030 is ripe for opportunities to grow and refine graduate education in the Marine Corps. Some of its stated goals are and should continue to be missions for the Marine Corps graduate education population. Therefore, the implications of this publication on the CEB are only now in their nascent stages.

2.4 Literature Review

What stood out from the relevant literature was not the presence of relevant literature, but rather the absence. It appears that researchers have greater incentives to find the predictors of performance during graduate education rather than after education and reentry into the workforce. For example, both Marks and Moss (2016) and Hall et al. (2017) focused their research on the performance of students within graduate education. On the other hand, Shultz and Zedeck (2011) focused their research on professionals after attending graduate education. However, there was a body of research on predictors for professional success

with the Marine Corps and Navy specifically.

Shultz and Zedeck (2011) focused their research on finding predictors of success for law school graduates working as lawyers professionally. Traditionally, law school admission has focused on undergraduate GPA and LSAT scores. They made the argument from their research that these factors should not weigh so heavily into the admissions calculus because they have limited predictive power when it comes to success as professional lawyers. Their predictors focused on “personality constructs, interests, values, and judgment” in addition to the traditional measures. Shultz and Zedeck (2011) differs from this research in that we are looking more at demographics and other more objective traits that can allow the Marine Corps to predict success in their 88xx billets. This study would be greatly improved by “personality constructs, interests, values, and judgement” (Shultz and Zedeck 2011) data; however, the Marine Corps currently does not collect this data for this study’s sample population.

Marks and Moss (2016) focused their research on predictors for success in law school. However, their findings may have bearings on this research. They found that GPA was overall a better predictor than LSAT scores. As it stands the Marine Corps does not require applicants to graduate education programs to take a standardized test. This practice is supported by Marks and Moss (2016). They also found that having a STEM undergraduate degree was a better predictor for success in law school than even degrees that are considered “pre-law.” If this applied to the Marine Corps, then the Corps should increase its demand for STEM undergraduates regardless of the graduate program to which Marines are applying. However, our findings differ from Marks and Moss (2016) in that STEM degrees had no influence on a Marine’s performance in an 88xx billet. The final finding that was of interest to this research was that work experience was also a good predictor of success. In their research, they concluded that “four to nine years of work is a sweet spot.” This validates the Marine Corps’s current practice of sending officers to graduate school at around the four to ten-year mark.

Hall et al. (2017) were looking for predictors of high-quality researchers. Ph.D. programs often focus their attention on the usual predictors for graduate school; however, the difference is the final goal. The final goal of graduate school is to produce expert practitioners in a specific field, whereas Ph.D. programs produce expert researchers in a particular field. Thus,

their research focused on predictors of student research productivity. They used the number of papers published as the categorical variable that they were predicting. They found that standardized testing was a poor predictor of performance, whereas letters of recommendation was a good predictor. Their recommendations were to de-emphasize traditional metrics like undergraduate GPA and standardized testing scores, and rather emphasize the holistic quality of the student. This research is different from ours in that they are still focusing on the quality of the student and not the quality of the professional.

Rateike (2017) influences this study, because he sought to find predictors that were influential in being selected by CCLEB for graduate education. He found that FITREP scores and The Basic School (TBS) rankings were very influential in predicting whether or not a Marine Officer was selected for graduate education. This study differs from his in that we are looking for predictors as to who should go to graduate school, whereas he was looking for predictors at who did go to graduate school. However, it may be of importance to this study to see if FITREP scores and TBS rankings are also good predictors of who the Marine Corps should send to graduate school.

The research done by Talaga (1994) is pertinent to this study because he was looking at the influence of graduate education on future career success. His research sought to find how graduate education influenced promotion to O-4, recommendations for early promotion, and early promotion. What he found was that graduate education was influential on the first of these three response variables. This research is different from ours in many regards. His population is Navy Surface Warfare officers, and ours is Marine Corps officers. His response variables involved promotions, and ours is FITREP scores. Finally, he was looking at the influence of graduate education, whereas we are looking at who should receive graduate education. Overall, his findings do not have much bearing on this study, but his research was loosely related.

The research done by Lianez et al. (2003) likewise looked at Marine Corps graduate education programs. Like Talaga (1994), they were looking to compare career outcomes between those that attended residential PME like EWS, non-PME like NPS, and those without any graduate level education. They found that PME was more of a positive factor than the other two categories. Their research was limited though, because they only had data for four years of FITREP scores. However, similar to this study they were looking to predict

FITREP scores. This research is different in that we are not using graduate education as a predictor for performance after graduate school. Rather we are looking at Marines prior to graduate school to see what traits indicate performance in their graduate school-specific billets.

Overall, this literature space is fairly open and undefined. The body of literature looking at predictors before graduate education for professional success after graduate education is very limited. Therefore, this study will seek to provide insights into this feature space. This is important for the Marine Corps as it invests into the education of its officer corps, and employs its graduate-educated officers in specific and critical billets.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 3: Methodology

In this chapter, we go through the process by which we clean, process, and manipulate the data. We speak to why some entries are omitted. We also speak about how we resolve the categorical predictors into factors and create the response variable. We cover our methodology for applying supervised machine learning, namely Regression Trees and Random Forest, to build a model from which we can make predictions.

3.1 The Data

The data set is from the MCTFS Data Warehouse (MCTFS Data Warehouse 2023) via the Manpower Information Systems Division. The sample population in question is all 88xx AMOS Marine Officers who acquired their 88xx AMOS from the years 2007 to 2021. 88xx Military Occupational Specialty (MOS)s are specialties that require a master's degree to fill. These master's degrees can be acquired either via ADP or SEP. This study uses the Marine-1 and Master Brief Sheet (MBS) data for the sample population. The Marine-1 data contains all the predictors. The MBS data has the Marines' FITREP data, which is our response variable.

The Marine-1 has the following categorical variables that are of interest to our research: PMOS, source of entry, Time in Service (TIS), undergraduate major subject (liberal arts or STEM), and education level (undergraduate or some graduate education). This data lacks any numeric variables which is atypical and significantly limits the effectiveness of both linear and machine learning models.

3.2 Data Cleaning and Processing

Using R (R Core Team 2023), we upload the MCTFS and MBS data from the MCTFS Data Warehouse (2023) into dataframes. From the former, we retain the predictors that are of interest to this study as well as Electronic Data Interchange Personal Identifier (EDIPI) data, which is used to merge the two dataframes. From the latter, we likewise retain the EDIPI as well as the Billet Military Occupational Specialty (BMOS), months for the reporting

period, and the three different scores for that particular FITREP. The three different FITREP scores are average-this-report, relative-value-at-processing, and relative-cumulative-value. The raw MCTFS data has 1177 entries, but this narrows down to 850 when merged with the MBS data and incomplete entries are dropped.

For the MCTFS dataframe many of the predictors have hundreds of factors; therefore, in order to effectively make a regression model we have to compress these factors down to a handful that can properly characterize the population. Additionally, some of the factor levels have only a couple of entries and thus are not useful to a model made up of fewer than a thousand rows.

The first predictor is the *civilian education certificate code*. This code indicates what level of education a Marine has achieved. For this study, all the empty rows are assumed to have a bachelor's. This is a reasonable assumption because it can safely be assumed if a Marine has attended graduate school he or she most likely had a bachelor's degree before doing so. Associates degrees are also modeled as bachelors. We model the one entry that has a doctorate as a master's. We remove the one row with a professional degree. None of the post-master's made it into the data set. We characterize those with only a high school diploma as having a bachelor's to improve the quality of the model because their numbers were so few. Therefore, we condensed this predictor's levels down to bachelor's degree and master's degree or undergraduate and graduate.

The next predictor is essentially identical to the previous one, and it is the *civilian education level code*. This code is an integer indicating the number of years of education that an individual has. The continuous variable is made into a factor with the levels of less than or equal to 16 years, "undergraduate," and greater than 16 years, "graduate." We will refer to this as "education level."

PMOS is our next predictor. This predictor has over a hundred levels, one for every MOS in the data set. We compress these down to "support," "ground," and "aviation." Support includes MOSs like logistics, supply, and other supporting establishment roles. MOSs like general officer were included in support. Ground includes what the Marine Corps colloquially calls combat arms MOSs. This includes infantry, tracks, tanks, artillery, etc. Aviation includes all pilot MOSs as well as other aviation support roles like aviation supply

or air traffic controller.

For the *current source of entry* factors, we reduce the levels to “Academy,” “PLC,” “OCC,” “ROTC,” and “other.” Academy includes all Naval, Military, and Air Force Academy graduates. PLC is all Marines who commissioned via Platoon Leader’s Course (PLC). OCC is all Marines who commissioned via Officer Candidates Course (OCC). ROTC is all Marines who commissioned via Reserve Officer Training Corps (ROTC). Finally, “other” includes all various and sundry commissioning methods not previously mentioned to include prior-enlisted commissioning routes. We will refer to this as “source of entry.”

For the *civilian education majors subjects code*, we use the levels of STEM and Liberal Arts. All degrees identified as Academy and any degree containing a moderate amount of math or science are folded into STEM. All others are considered Liberal Arts. We will refer to this as “major subject.”

We group the predictor *years of service* into three categories: fewer than five years, five to nine years, and greater than nine years. We will refer to this as TIS.

We build our models using the five categorical variables of education level, PMOS, source of entry, major subject, and TIS. We have a total number of 850 data entries.

We also look at marital status, race, gender, and the number of dependents. We retain them in the data-set to gain insight; however, we do not use them to make predictions because the Marine Corps cannot make assignment decisions based on these characteristics. We compress marital status to the factor levels of single, married, and divorced. We characterize the one widowed data entry in the data set as “married.” We compress race codes to the levels of African-American, Asian-American, Caucasian, and Other. We simplify the number of dependents to 0, 1, 2, 3, and greater than 3.

3.3 Response Variable

To select a response variable, we have to find a variable that generally characterizes the performance of the Marine in the role for which they attended graduate school. The logical variable for this is the average score from a Marine’s FITREP. So we pull the average scores from the FITREPs associated with 88xx billets. 88xx billets are billets that are assigned

exclusively to Marines with graduate degrees in which they will use the graduate degree. However, within these billets, an individual Marine will often have multiple FITREPs depending on how long they are in the billet, whether their RS changes, or if they have a second tour as an 88xx. Therefore, our response variable is a time-weighted FITREP score. For each Marine in the sample, let

$$response = \frac{\sum_{i=1}^T f_i t_i}{\sum_{i=1}^T t_i},$$

where f_i is the score from FITREP i , t_i is the duration in months of the period of observation for FITREP i , and T is the total number of reports for the given Marine. This yields one time-weighted response value per Marine.

For example, a Marine filling an 88xx billet may have multiple FITREPs as an 88xx. The reason for this is that a Marine's reporting time as an 88xx may be divided into multiple FITREPs for any number of reasons. These reasons include multiple tours as an 88xx, change of rank, change of RS, annual requirements, etc. Thus, if a Marine has a FITREP where his observation time is 12 months and a FITREP where his observation time is 6 months, the former will have twice the weight in the response variable than the latter. The end result is all of a Marine's FITREP scores are condensed into one response variable.

This equation includes a number of assumptions. First, a RS's given average this report scores do not change over time. Generally, this is not true. As a RS reviews more Marines, the relative values of the Marines in his profile tend to decrease over time. However, this is a reasonable assumption for this study because we are not comparing Marines between calendar years. For instance, we are not comparing Marines with FITREPs from 2010 to Marines with FITREPs in 2020. Rather, we are looking for attributes that indicate performance across the entire population regardless of year. The second assumption is that the time-weighted averages are a more accurate indicator of a Marine's performance than averages. We believe this is a valid assumption because there is a learning curve for every billet. If a Marine spends a short amount of time in a billet, then he will not properly learn the job and so his relative value score may reflect that detail. Whereas, if a Marine has an entire year in a specific billet, then his performance should be greater than that of the short-timer. Our response variable reflects this learning curve. A third assumption we

make is that FITREP scores are not rank skewed. This means that we are assuming that if a major has a higher response variable than a captain it is because of performance and not rank. This assumption will reduce the strength of any findings where greater TIS leads to better FITREPs, because it is generally accepted that the higher the rank the better average FITREP score.

For this study, we compute the response variables for three different FITREP scores. The first is the “average-this-report.” This score is the average of all the assigned scores, with a minimum of 1 and a maximum of 7. This score remains the same throughout a Marine’s career. The second score we compute is the “relative-value-at-processing.” This score takes into account a RS profile and adjusts the score depending on whether the RS ranks the given Marine’s FITREPs low or high relative to other Marines he or she has evaluated for that grade. This relative value is computed when the FITREP is submitted and likewise does not change for the rest of the Marine’s career. A relative value is a score that ranges from 80, a RS’s lowest awarded FITREP, to 100, a RS’s highest awarded FITREP. Not all Marines will have a relative-value-at-processing if their RS does not have a profile. A RS does not have a profile until they have written three FITREPs for that particular grade. We have 435 Marines with relative-value-at-processing scores compared with our larger population of 850 for whom we have average-this-report scores. The third score we compute is “relative-cumulative-value.” As a RS writes more FITREPs his or her profile may change, thus the relative values of the FITREPs in their profiles change. The relative-cumulative-value is the relative value of the FITREP when the data is exported from Total Force Data Warehouse (TFDW). As with the relative-value-at-processing, only Marines whose RSs have profiles will have this score. Thus, we have 637 entries with relative-cumulative-value scores out of the 850.

3.4 Regression Trees

For this study, we implement regression trees (James et al. 2017) using the *rpart* (Therneau and Atkinson 2022) function in R (R Core Team 2023). This algorithm takes the split and computes the means and the deviances for each split. It seeks to find the split or splits that decrease the deviance. The algorithm will find the split where the decrease is largest. The algorithm stops splitting when a split is too small to logically continue to make splits. For

example, we do not want a split that has a single or even a handful of data entries. For predictions, the algorithm uses these splits to classify the inputs and predicts for each input the average response found for that split.

3.5 Random Forests

A Random Forest model is an example of an Ensemble Model (James et al. 2017). An ensemble is a combination of other models. Random Forests are models that contain multiple regression tree models. The algorithm takes the training data and creates a tree. Unlike with *rpart* (Therneau and Atkinson 2022), it does not try to make the best tree. Rather it perturbs the data little by little making new trees that are marginally better than a naive model. A Random Forest model generates multiple regression trees by bootstrapping and by sampling the subset of predictors. From here these aggregated trees are used to make more flexible predictions that have not been over-fitted. The goal of the Random Forest is a model that has both low bias, and low variance.

For this study, we use three different types of Random Forest models. The models are: Random Forest (Liaw and Wiener 2002), Bagging (Bootstrap Aggregating) (Hothorn and Zeileis 2015), and *cforest* Random Forest (Hothorn and Zeileis 2015). All these models employ the bootstrap and perturbations of the data to find the best set of trees from which to make predictions. For all these models, we use all the available predictors. The difference between *randomForest* and *cforest* is that *randomForest* uses unweighted averages and *cforest* uses weighted averages (Shuvayan 2015). We use multiple models so as to find the model which best predicts our response variable.

3.6 Support Vector Machine (SVM)

SVM (James et al. 2017) seeks to select the best hyperplane through the data in order to conduct binary classification. SVM uses optimization to find the best hyperplane that separates the two data categories. It does this by finding the distance to a proposed separator. SVM is specifically used in this study to confirm whether or not major subject is an influential factor in determining the response variable. Therefore, we adjust the data to use the response variable and the other predictors to classify whether or not someone had a liberal arts degree

or a STEM degree. We seek further depth in this particular area, because it has implications for the CEB.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 4: Analysis

This chapter starts with our findings from our Exploratory Data Analysis (EDA) on the MCTFS data. (MCTFS Data Warehouse 2023) From there we discuss the findings from our machine-learning regression models. Finally, we speak to our supplemental analysis that we conduct to confirm various insights gleaned from the regression modeling.

4.1 EDA

4.1.1 Response Boxplots

With our EDA, we start by observing the boxplots of FITREP average-this-report scores for our five predictors, see Figure 4.1. SOE is the source of entry. As seen from the SOE plot, the inter-quartile distance between the first and third quartiles are all about equal across the categories. The median for the Naval Academy is a little higher than the rest and that for PLC is a little lower. However, overall there is very little variance among the categories. This means that the source of entry probably has very little predictive power.

The next boxplot “YRS COMM SER” is TIS. Generally, the longer the service time the higher the response variable. Essentially, Majors get higher FITREP scores than Captains when occupying 88xx billets. However, even this varied only marginally. Additionally, this variance disappears when looking at the relative-value-at-processing and relative-cumulative-value response variables.

The comparison between the boxplots of the PMOS yields even fewer insights than the former two. The ranges and the means are all generally the same. This is surprising because we hypothesize that ground MOSs on average have better FITREPs due to the Marine Corps directing many of its more talented Marines to the ground combat MOSs. However, the data does not seem to indicate the general MOS community makes a difference.

The one predictor with a fair amount of variance across the categories is “YEARS EDUCATION” or education level. Those with some level of graduate education receive better

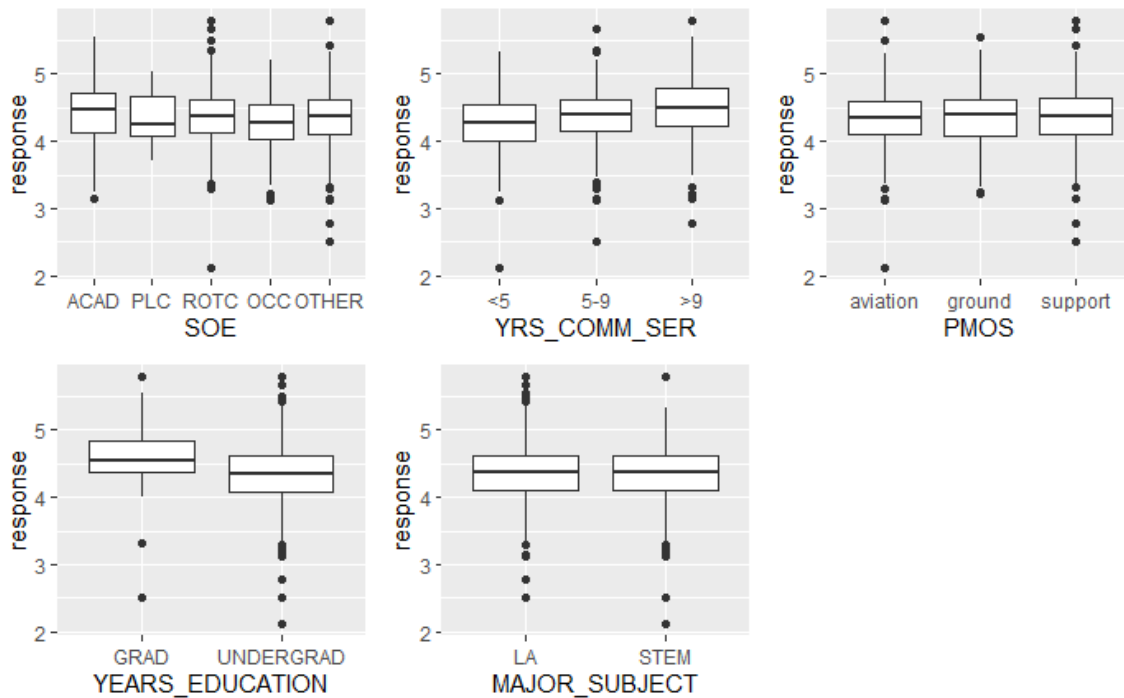


Figure 4.1. Contains the boxplots of all the categorical variables employed in the model to make predictions. The response is the FITREP average-this-report variable created by this study. Top left is the source of entry which was either the Service Academies (ACAD), PLC, ROTC, OCC, and other. Top middle is the number of years of commissioned service binned into less than 5 years, 5 to 9 years, and greater than 9 years. Top right is PMOS binned into aviation, ground, or support. Bottom left is education level binned into graduate and undergraduate. Bottom right is the major subject binned into liberal arts (LA) or STEM. Data adapted from MCTFS Data Warehouse (2023).

FITREPs than their peers who have no graduate education upon attending their particular Graduate Education Program (GEP). This is a limited conclusion due to the sample size. There are only 49 individuals with some level of graduate education while the remaining 801 have none. This disparity in sample sizes greatly reduces the predictive power of this predictor.

The last predictor is the major subject. We hypothesize that those with STEM degrees perform better than their peers with liberal arts degrees. However, the boxplots seem to indicate there is no difference between the two categories. This is robust because the sample

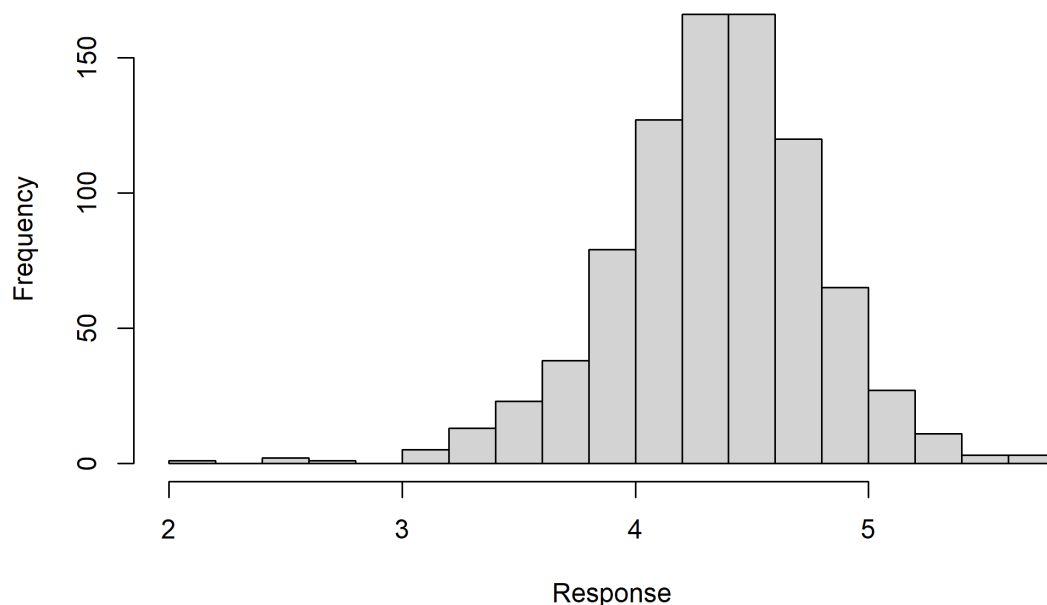


Figure 4.2. The histogram plot of the response variable of FITREP average-this-report. Data adapted from MCTFS Data Warehouse (2023).

sizes between STEM and liberal arts are both of moderate size, with liberal arts at 535 and STEM at 315. This is a key insight for the CEB because they try to find Marines with STEM undergraduate degrees to send to school to fill 88xx billets. However, this data indicates that a STEM degree is not necessary at all. Now, this insight does not apply to STEM versus liberal arts Marines’ performance at school. It is strictly limited to their performance in filling 88xx billets. Later in this chapter, we delve deeper into this subject with our SVM model.

4.1.2 Normality Testing

In Figure 4.2 we plot the histogram of the response variable. As seen from the histogram the response variable is generally Gaussian with a heavy left-side tail.

In order to test for normality in the response error, we build a multivariate linear model predicting the FITREP average-this-report response variable. From there we plot our model, see Figure 4.3. As you can see in the figure the mean is generally about zero and the variance

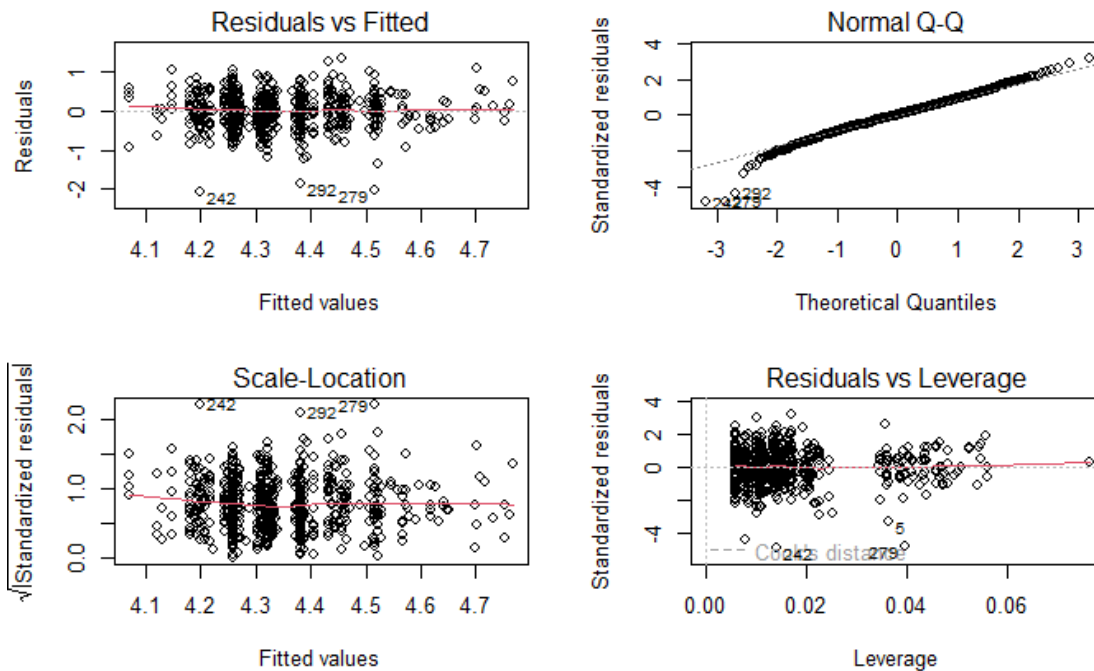


Figure 4.3. The plots of the linear model residuals, Q-Q plot, scale-location, and the Leverage Residuals. The residuals versus fitted plot shows that the errors are roughly homoscedastic. The Q-Q plot shows that the data is roughly normal. The scale-location similarly points to the average error being zero. Finally, the leverage residuals plot shows that there are a significant subset of points that have high leverage. Data adapted from MCTFS Data Warehouse (2023).

is fairly constant. The Q-Q plot generally confirms that the data is Gaussian. However, there are some oddities in the data. As seen from the leverage plot, a whole subset of points have a leverage greater than 0.04, whereas the preponderance of the data lies at or below 0.02. This could mean that making predictions from a linear model might be quite difficult. This is further exacerbated by the lack of any numeric predictors. They are all categorical.

4.2 Supervised Machine Learning Models

In this section, we report the results found from the machine learning models. We start with the *rpart* regression trees, and then move to the various ensemble models that we make. The ensemble models include two Random Forest models and a bagging model.

RMSE Table	Response	Response.RAP	Response.RC
Decision Tree	0.036	0.638	0.387
RandomForest	0.033	0.628	0.322
Bagging	0.036	0.568	0.339
cforest	0.049	0.568	0.242
Linear Model	0.035	0.556	0.388
# of Data Points	850	435	637

Table 4.1. The table of the Root Mean Squared Error (RMSE) values of the linear model and machine learning models for the three different response variables. These values are measured from the FITREP response variable which have a minimum of 1, a maximum of 7, and an average around 4.4.

4.2.1 Regression Trees

For our average-this-report response variable, we build a tree that can be insightful for analysis, as seen in Figure 4.4. The first split made by the algorithm is at TIS. Those with greater than 9 years receive better FITREPs with an average of 4.51 and those with less than 9 years have an average of 4.30. This generally means that majors selected by the CPIB have better scores than their captain colleagues who are selected by the CCLEB. This could be a result of the favorable impression higher rank has in the Marine Corps or it could be resultant of the maturation process of Marine Corps officership. Within the fewer than 10 years of commissioned service branch there is an additional split between officers who have fewer than five years TIS at 4.25 and those with more at 4.35. This might indicate more seasoned officers make better 88xx. Within the greater than nine years branch, the tree first splits at source of entry being OCC as being lower with a score of 4.18 compared to 4.57. It then splits the higher group again at undergraduate education with a score of 4.52 compared to those with some graduate education at 4.86. These splits seem to indicate that officers commissioning via OCC perform worse than their peers, and that some level of previous graduate education is desirable as an 88xx. However, these conclusions are not very strong, because the OCC leaf contains 3.2% of the population and the graduate education leaf only 2.4% of the population. The only moderately firm conclusion from this model is that greater TIS generally means better FITREPs. The predicted versus actual values can be seen in Figure 4.5. The data does not strongly group together around the regression tree predictions

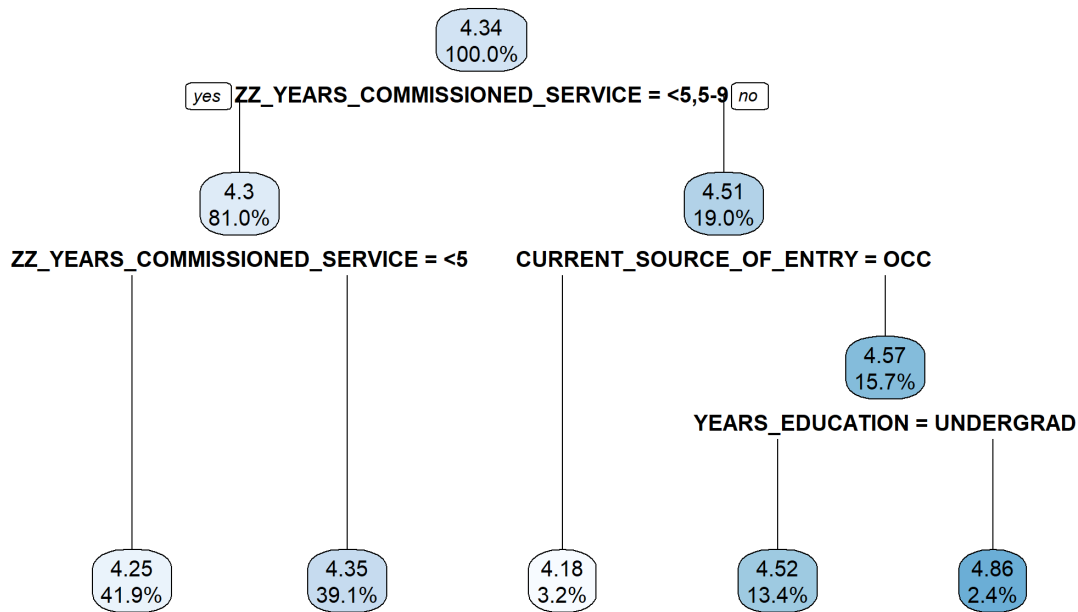


Figure 4.4. The *rpart* decision tree for the average-this-report response variable broken down into five groups with the predicted response value and the percent of the population in that group. If a decision is true then data goes left. The shading qualitatively reflects the average in each node.

suggesting any conclusions drawn from this model are not strong.

The relative-value-at-processing response variable regression tree model is not quite as insightful as the regular model mentioned above. In Figure 4.6, we see that the algorithm first made a split on PMOS being an aviation MOS. Those with aviation PMOSs are around a 92 in their RS's profile, whereas those with support or combat arms PMOSs are around a 94. The right branch then forks at source of entry being OCC and other with a relative value of 93, whereas their counterparts from other commissioning sources average around a 94. These differences in the world of FITREPs are not significant enough to justify concluding that one subset of Marines makes for better 88xx Marines than another. A model that makes three predicted values is too simplistic a model to provide much real insight into the data. Additionally, this model is further weakened by containing around half of the original sample size used in the average-this-report model.

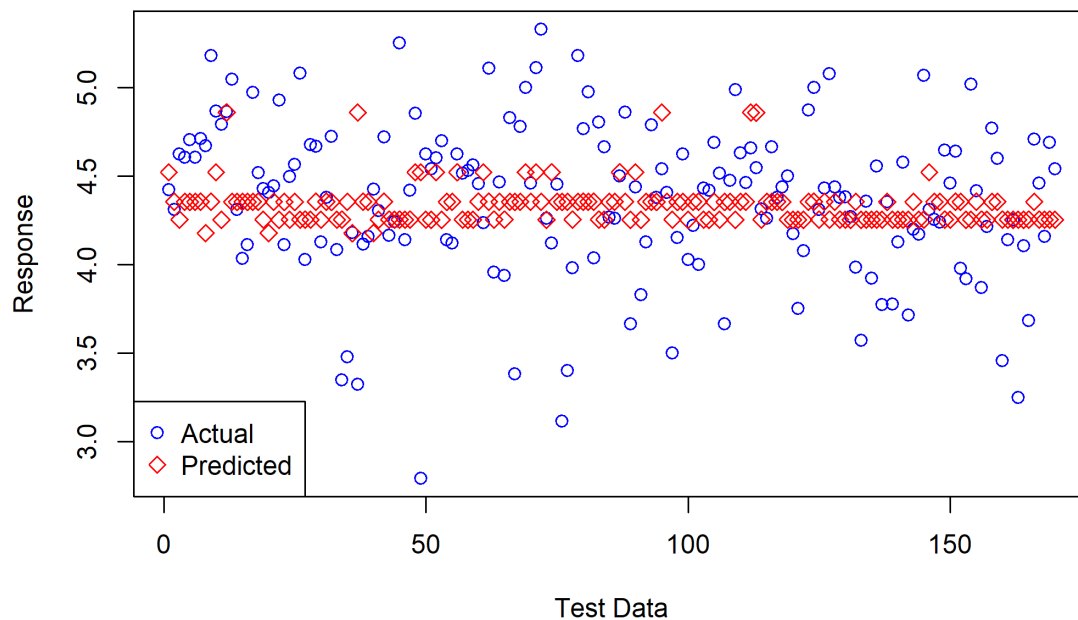


Figure 4.5. The comparison plot of the average-this-report response variable between where the *rpart* tree predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.

The relative-cumulative-value response variable regression tree model offers even more limited conclusions than the relative-value-at-processing model. In Figure 4.7, we see that the algorithm split at only PMOS of aviation with a relative value of 91 against their counterparts in the support and combat arms MOSs with a relative value of around 93. These differences are again quite subtle, and perhaps the board should lean away from sending aviation PMOSs to graduate education, but that conclusion is not strongly supported. Again the simplicity of this model makes it lack predictive power. Also this model contained about three quarters of the overall dataset, which degrades the value of its insights.

Overall, it is expected that the results from a single tree are not particularly powerful. A regression tree lacks the flexibility to model complex data, especially data with as small a number of predictors as this dataset, as well as containing exclusively categorical predictors. Perhaps one conclusion that can be drawn from these regression trees is that as it stands

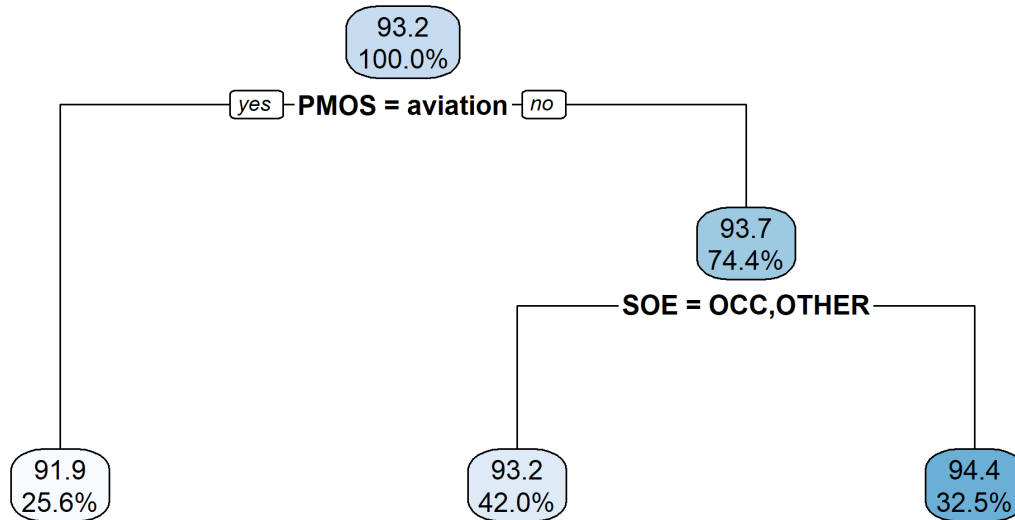


Figure 4.6. The *rpart* regression tree for the relative-value-at-processing response variable broken down into five groups with the predicted response value and the percent of the population in that group. If a decision is true then data goes left. The shading qualitatively reflects the average in each node.

this dataset does not contain the predictors necessary to determine substantial differences among subsets of Marines. This would mean the board should not make selection decisions based on major subject, PMOS, and source of entry. However, there is a small case to be made that the board could make determinations based on TIS as well as education level. These conclusions though are tenuous at best with only the regression tree models.

4.2.2 Ensemble Models

For this section we discuss our *randomForest*, bagging, and *cforest* ensemble models. Each model uses all 850 observations and all five predictors.

The first ensemble model we build is with the *randomForest* algorithm. We create 1000 trees total for this model and found our RMSE increased if we did any more than 1000.

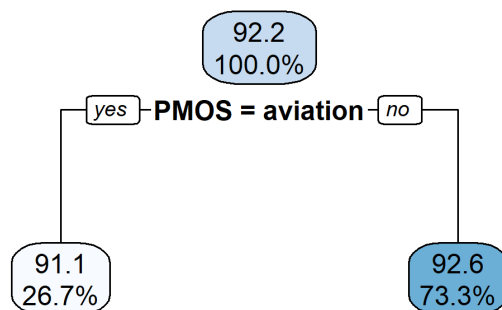


Figure 4.7. The *rpart* regression tree for the relative-cumulative-value response variable broken down into five groups with the predicted response value and the percent of the population in that group. If a decision is true then data goes left. The shading qualitatively reflects the average in each node.

We have the algorithm randomly sample one predictor out of the five predictors at each split because Hastie et al. (2009) recommends the default number of predictors to choose is $p/3$ where p is the number of all predictors. We do not vary this to other settings. The actual versus predicted values can be seen in Figure 4.8. As seen from this figure, the model does not fit to the data well. Its accuracy appears to come from having modeled the mean quite thoroughly. However, this model has the smallest RMSE of any of our models for the average-this-report response variable as seen in Table 4.1. Our *randomForest* models for the relative-value-at-processing and relative-cumulative-value response variables are a full order of magnitude larger in their RMSEs thus indicating their lack of predictive power. Therefore, for any insights and predictions, we refer to the average-this-report response variable model.

The second ensemble model uses the bagging algorithm to make a random forest model.

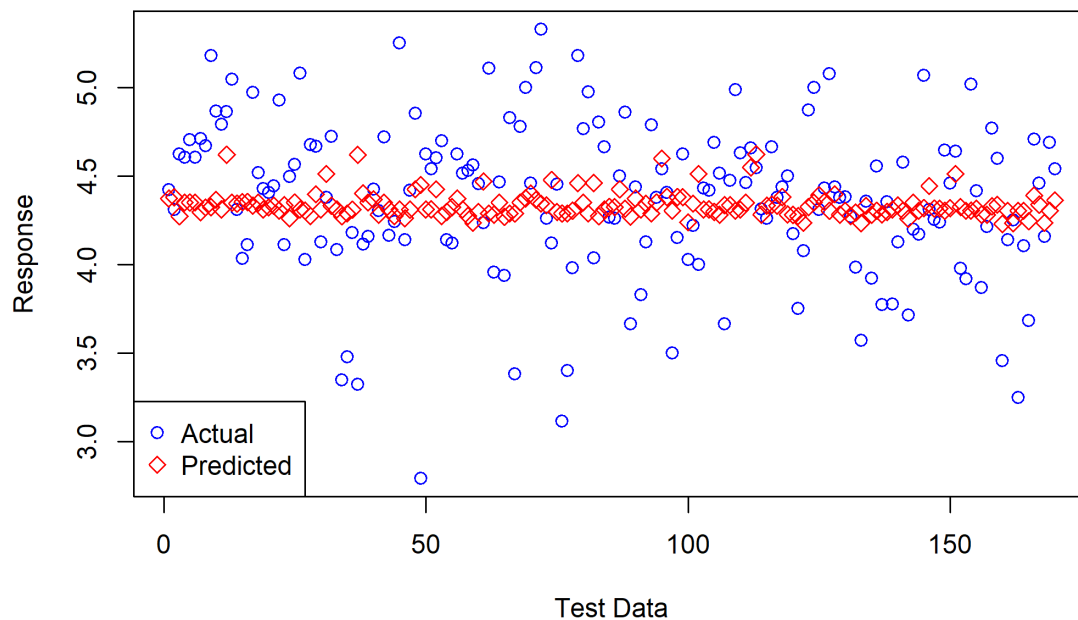


Figure 4.8. The comparison plot of the average-this-report response variable between where the randomForest algorithm predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.

For this model, we grow 500 trees. The actual versus the predicted values can be seen in Figure 4.9. This model does not appear to have fit the data well with only a handful of predictions varying significantly from the predicted mean. This model has the second smallest RMSE of our ensemble models, but its RMSE is larger than the linear model and equivalent to the regression tree model for the average-this-report decision variable. As with the *randomForest* model, this algorithm also has an order of magnitude larger RMSE for the relative-value-at-processing and relative-cumulative-value response variables. Thus, it is not used for any predictions or insights.

The final ensemble model we build is the *cforest* random forest model. For this model we grow 500 trees. The actual versus the predicted values can be seen in Figure 4.10. As seen from this figure, it appears the model tried to follow the data, more so than the previous algorithms. However, it generally fails to do this because its RMSE is larger than all our

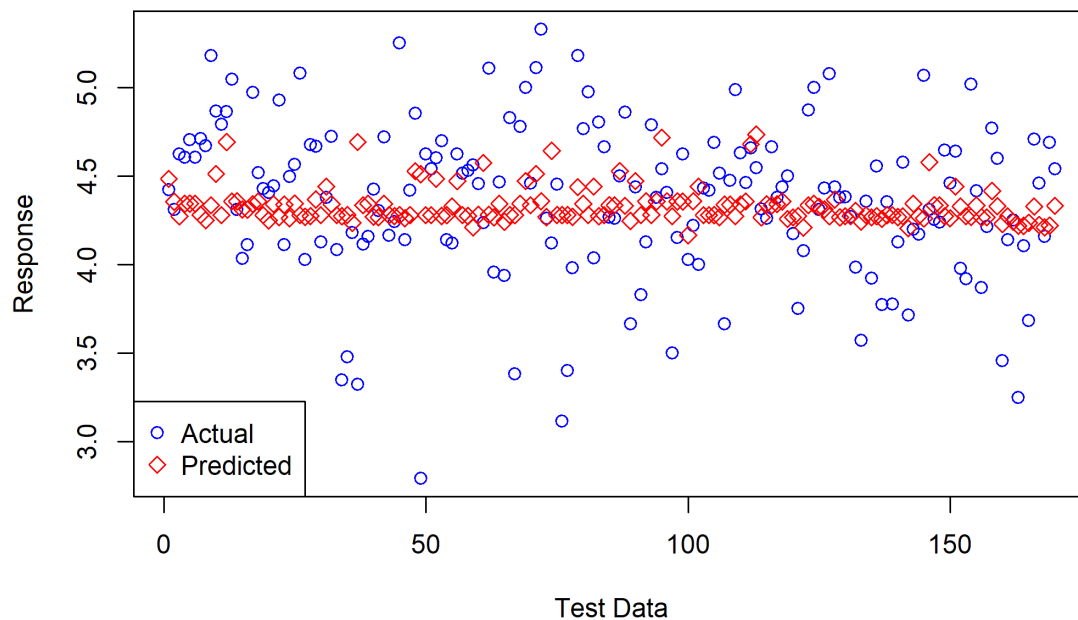


Figure 4.9. The comparison plot of the average-this-report response variable between where the bagging algorithm predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.

other models. Additionally, its relative-value-at-processing and relative-cumulative-value response variable models are similarly powerless in prediction.

As seen from the scatterplots, it appears that with the data available the ensemble models struggled to accurately model the data. The response appears to be ambivalent to the predictors provided. This would suggest that the predictors studied are not important for determining the success or failure of a given Marine performing in 88xx billets.

4.3 Linear Model

For building a linear model, we use the *stepAIC* (Venables and Ripley 2002) function to find an effective model from the five predictors provided. The final model predictions can be seen in Figure 4.11. The linear model attempted has the second best RMSE following behind only the *randomForest* ensemble model. The coefficients of the final model are in

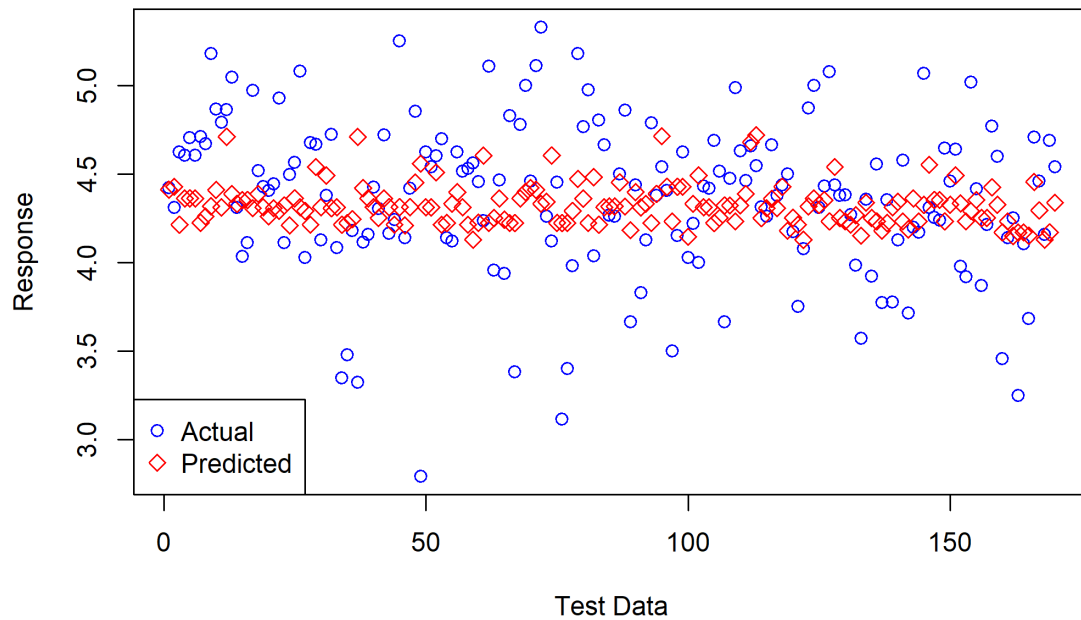


Figure 4.10. The comparison plot of the average-this-report response variable between where the *cforest* algorithm predicted a response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.

LM Table	Intercept	SOE(OCC)	SOE(other)	TIS (5-9)	TIS (>9)	UGRAD	STEM
Coefficient	4.50	-0.14	-0.06	0.11	0.25	-0.19	-0.06

Table 4.2. The table of the coefficient values of the linear model created using *stepAIC*.

Table 4.2.

All the variables are binary. OCC and OTHER refer to sources of entry OCC and other. TIS (5-9) and TIS (>9) are both TIS variables. ‘UGRAD’ is the education level of undergraduate. Finally, STEM refers to those possessing STEM undergraduate degrees as their major subject.

The final model produced from *stepAIC* has small negative coefficients for source of entry

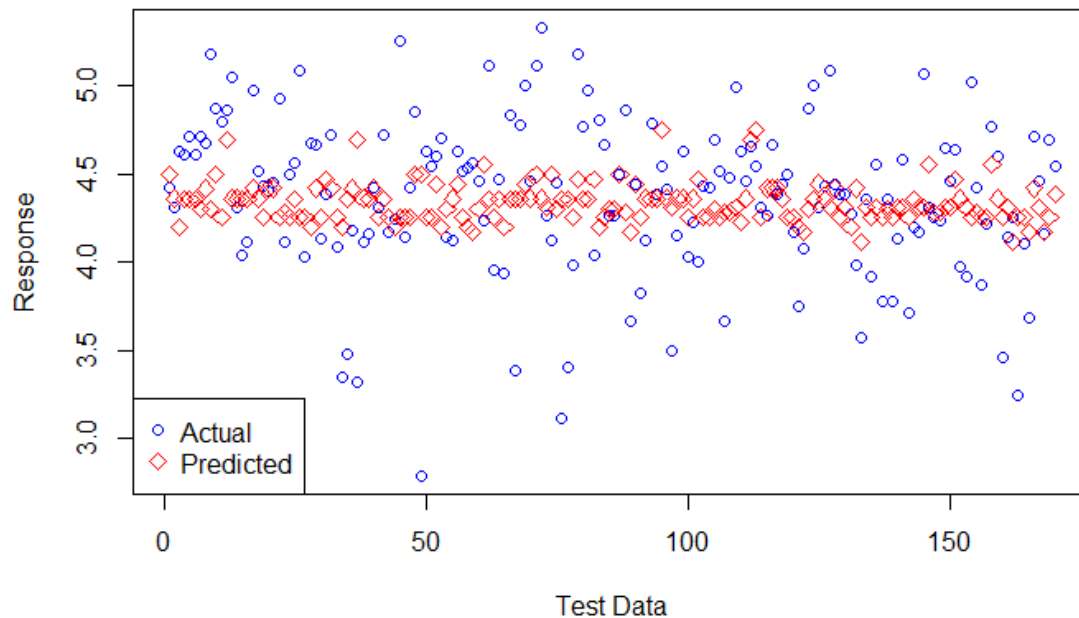


Figure 4.11. The comparison plot of the average-this-report response variable between where the linear model's predicted response value for a given data point versus its actual response. The response is the average-this-report response variable created by this study. The x-axis is the observation number.

“other” and major subject “STEM.” This means that those commissioned from some alternative commissioning source or have STEM degrees generally perform slightly worse than their colleagues. However, these are on the order of -0.06 which for a score that is generally around 4 to 5 is inconsequential. The model has larger negative coefficients (around -0.14 to -0.19) for source of entry OCC and education level of undergraduate. Thus, those from OCC and those without previous graduate education generally perform slightly worse than their colleagues. However, these numbers are again quite small. Finally, the model has a modest positive coefficient for TIS of 5-9 (around 0.11), and a larger one for TIS greater than 9 (around 0.25). These numbers indicate that greater TIS leads to greater performance as an 88xx. Again, these numbers are quite small, and thus are not particularly influential relative to the intercept which is 4.5.

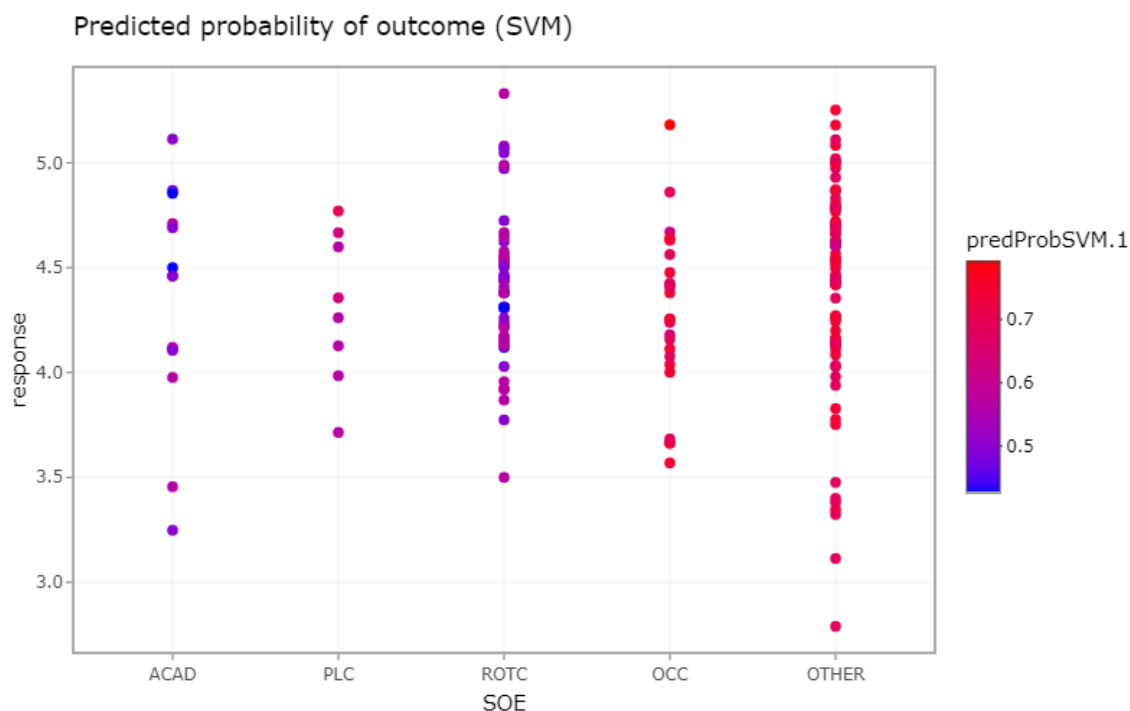


Figure 4.12. The plot of the predicted probabilities in color scale with the source of entry on the x-axis and the response variable on the y-axis. The response here is the average-this-report FITREP variable created by this study.

4.4 SVM Model of Major Subjects

Our SVM model to investigate the influence of a Marine’s undergraduate major subject reinforces our findings from our regression tree, ensemble, and linear models. The predicted probability of outcomes can be seen in Figure 4.12. The plot shows that the means of each of the five sources of entry hover around the mean (4.37). Additionally, the colors of the predicted outcomes seem to likewise hover around the mean. If STEM specifically were correlated with a higher FITREP score then we would see more blue and less red above the mean. This is not the case. Rather what we see is the blue versus red balance seems to correlate to the source of entry. OCC and “other” sources of entry are predominantly liberal arts, whereas the Academy and ROTC appear to be predominantly STEM. PLC appears to be somewhere in the middle. This means that the SVM used the source of entry and not the response variable to construct its hyperplane. Also the model’s classification accuracy is very low, around 66%, only 7% better than the naive model, making its insights into the

data particularly limited. Overall, this further reinforces that there is very little correlation between STEM and FITREP scores for Marines filling 88xx billets.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5: Conclusions and Recommendations

This chapter describes some of the limitations of our approach, some conclusions that can be drawn from the data, recommendations for the CEB, and finally, some recommended further research.

5.1 Limitations

Overall, in this study, we did not find any predictors that are significantly influential in predicting the success of 88xx in their utilization tours. We hypothesized that STEM would be an indicator of performance over Marines with liberal arts degrees. We also hypothesized that source of entry may perhaps have some influence due to the heavy emphasis on technical degrees for those graduating from the Academies. Finally, we hypothesized that Marines from combat arms PMOSs would perform better than their peers due to the Marine Corps frequently sending in TBS their most talented Marines to ground combat MOSs. The data simply did not support these hypotheses.

For our findings, there is actually a slight negative correlation between STEM and our response variable; however, it is not significant enough to suggest the reverse conclusion. Additionally, Academy Marines appear to perform comparably to their peers. Finally, there is insufficient evidence to suggest ground combat PMOSs perform better than their aviation and support PMOS peers. Therefore, the board's current practice of diligently searching for Marines with technical undergraduate degrees, drawing heavily from the Academies, is an unnecessary restraint on their selection process. We also recommend no vetting be done based on PMOS for the purposes of trying to find the highest-performing 88xx.

5.2 Conclusions

Our other two predictors, TIS and education level provide us with some limited conclusions. Generally, the more experienced a Marine is upon gaining an 88xx BMOS the higher his performance. There is a slightly increasing trend as we go from fewer than five years to five to nine years and then to greater than nine years. This could perhaps be explained by majors

receiving better FITREPs than their captain colleagues based on a rank skew perception of performance, or it could be a function of greater maturity as a Marine officer. The relative-value-at-processing and relative-cumulative-value response variables showed very little variability between the three categories, thus undermining this conclusion. Therefore, there is insufficient evidence to recommend any changes to the board on account of TIS.

The level of previous education before graduate school seems to indicate a positive correlation between having some level of graduate education before attending graduate school. This conclusion is limited in its strength due to the few Marines possessing some graduate experience. The subset containing undergraduate degrees or less was significantly larger thus limiting a model's ability to use this predictor to predict. Perhaps, with all things being equal the board could opt for a Marine with some graduate education over the one with none.

5.3 Recommendations for the CEB

We recommend the board discontinue looking for Marines with technical undergraduate degrees or some experience with math and science. This practice indubitably reduces the size of the population from which they can select. This overall will limit the quality of the candidates they are able to select. The graduate programs appear to be providing the skills necessary to succeed in an 88xx billet regardless of whether a Marine previously had technical education. Now, this conclusion does not apply to performance in school. This practice may lead to a decrease in performance in school; however, this is not studied here. Additionally, a Marine's graduate GPA is of little importance to the Marine Corps. The Marine Corps is more concerned with a Marine's performance in their assigned job.

We recommend the board gives preference to Marines with some level of graduate education prior to attending ADP or SEP. We would not place heavy emphasis on this feature but rather use it to break ties.

We recommend the board continues to primarily make determinations based on a Marine's desirability for education and their competitiveness against their peers. As it stands the data available has not given the means to predict a Marine's competency for a given 88xx billet assigned to their specific graduate education program.

Finally, we recommend the Marine Corps draw a more comprehensive data picture of Marines sent to graduate school to better understand this population.

5.4 Further Research

We recommend acquiring TBS data to see if there are any predictors from that dataset that could be used to predict performance as an 88xx. We recommend creating a survey study to see if any survey questions with their responses can be used to more accurately predict 88xx performance. This survey can include anything from hobbies to interest in desk jobs to computer competency. The limitation of such a survey will be the imagination of the surveyor.

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

- Berger D (2021) *Talent Management 2030*. TM2030, United States Marine Corps, <https://www.marines.mil/Talent-Management-2030/>.
- Foley D (2022) NPS-23-M056: Refining the Selection Process of the Commandant's Education Board (CEB), <https://nps.edu/web/naval-research-program>.
- Hall JD, O'Connell AB, Cook JG (2017) Predictors of Student Productivity in Biomedical Graduate School Applications. *PLoS ONE* 12(1) (January), <https://doi.org/10.1371/JOURNAL.PONE.0169121>.
- Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, <https://doi.org/10.1007/978-0-387-84858-7>.
- Hothorn T, Zeileis A (2015) partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research* 16:3905–3909, <https://jmlr.org/papers/v16/hothorn15a.html>.
- James G, Witten D, Hastie T, Tibshirani R (2017) *An Introduction to Statistical Learning with Applications in R*, <https://doi.org/10.1007/978-1-4614-7138-7>.
- Lianez R, Zamarripa LR, Kocher KM (2003) The effects of U.S. Marine Corps officer graduate education programs on officer performance: A comparative analysis of professional military education and graduate education. M.A. thesis, Dept. of Natl. Sec. Aff., NPS, Monterey, CA, USA, <https://calhoun.nps.edu/handle/10945/1092>.
- Liaw A, Wiener M (2002) Classification and regression by randomforest. *R News* 2(3):18–22, <https://CRAN.R-project.org/doc/Rnews/>.
- Marks AB, Moss SA (2016) What Predicts Law Student Success? A Longitudinal Study Correlating Law Student Applicant Data and Law School Outcomes. *Journal of Empirical Legal Studies* 13(2) (June), <https://doi.org/10.1111/jels.12114>.
- MCTFS Data Warehouse (2023) Data provided to the author via website download, 25 January.

- R Core Team (2023) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, <https://www.R-project.org>.
- Rateike CN (2017) Quantitative analysis of high-quality officer selection by Commandant's Career-Level Education Board. M.A. thesis, Dept. of Natl. Sec. Aff., NPS, Monterey, CA, USA, <https://calhoun.nps.edu/handle/10945/53036>.
- Shultz MM, Zedeck S (2011) Predicting Lawyer Effectiveness: Broadening the Basis for Law School Admission Decisions. *Law & social inquiry* 36(3) (June), <https://doi.org/10.1111/j.1747-4469.2011.01245.x>.
- Shuvayan (2015) Difference between cforest and Random Forest in R. Analytics Vidhya. Accessed May 3, 2023, <https://discuss.analyticsvidhya.com/t/difference-between-cforest-and-random-forest-in-r/5455>.
- Talaga MT (1994) An econometric analysis of the effect of fully-funded graduate education on performance for Surface Warfare officers. M.A. thesis, Dept. of Natl. Sec. Aff., NPS, Monterey, CA, USA, <https://calhoun.nps.edu/handle/10945/28136>.
- Therneau T, Atkinson B (2022) *rpart: Recursive Partitioning and Regression Trees*, <https://CRAN.R-project.org/package=rpart>, r package version 4.1.19.
- Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*, 4th ed. (Springer, New York), <https://www.stats.ox.ac.uk/pub/MASS4/>, ISBN 0-387-95457-0.

Initial Distribution List

1. Defense Technical Information Center
Ft. Belvoir, Virginia
2. Dudley Knox Library
Naval Postgraduate School
Monterey, California



DUDLEY KNOX LIBRARY

NAVAL POSTGRADUATE SCHOOL

WWW.NPS.EDU

WHERE SCIENCE MEETS THE ART OF WARFARE