NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

IMPACT OF JUNIOR RESERVE OFFICERS' TRAINING CORPS ON THE QUANTITY OF ENLISTMENTS

by

John H. Panzer

June 2020

Thesis Advisor:	Samuel E. Buttrey
Second Reader:	Sean Clement, TRAC-Monterey

THIS PAGE INTENTIONALLY LEFT BLANK

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

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

4. TITLE AND SUBTITLE
IMPACT OF JUNIOR RESERVE OFFICERS' TRAINING CORPS ON THE QUANTITY OF ENLISTMENTS

5. FUNDING NUMBERS

6. AUTHOR(S) John H. Panzer

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 Junior Reserve Officers' Training Corps mission is to improve the citizenship of its students, and it achieves its mission by providing leadership and military skills to our nation's youth. However, the Army wants to create a more holistic approach to determine additional benefits that the JROTC program may provide. To meet this requirement, we investigate a potential ancillary benefit in understanding the JROTC program's effect on the quantity of enlistments. We collect demographic, economic, geographic, enlistment, and other data from primarily publicly available United States government agencies at the ZIP Code Tabulation Areas level. We focus our analysis using multiple regression techniques to investigate the potential impacts of a JROTC program on enlistments in ZIP codes, or ZCTAs, and their surrounding areas. We find that there is a positive association between the proximity to a JROTC program and the number of enlistments; however, the larger effects occur in the interactions among the JROTC distance factor, veteran population percentage, and 18- to 24-year-old population size. The combination of higher veteran rates and having a JROTC within a ZIP code is considerably more impactful than the individual factor benefits alone. Additionally, we predict the net improvement associated with establishing a new JROTC in ZIP codes to identify the ZIP codes that are best in terms of increasing the number of enlistments.

14. SUBJECT TERMS
JROTC, Junior Reserve Officers' Training Corps

15. NUMBER OF PAGES
71

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

IMPACT OF JUNIOR RESERVE OFFICERS' TRAINING CORPS ON THE QUANTITY OF ENLISTMENTS

John H. Panzer
Captain, United States Army
BBA, Florida Atlantic University, 2005

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2020

Approved by: Samuel E. Buttrey
Advisor

Sean Clement
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The Junior Reserve Officers' Training Corps mission is to improve the citizenship of its students, and it achieves its mission by providing leadership and military skills to our nation's youth. However, the Army wants to create a more holistic approach to determine additional benefits that the JROTC program may provide. To meet this requirement, we investigate a potential ancillary benefit in understanding the JROTC program's effect on the quantity of enlistments. We collect demographic, economic, geographic, enlistment, and other data from primarily publicly available United States government agencies at the ZIP Code Tabulation Areas level. We focus our analysis using multiple regression techniques to investigate the potential impacts of a JROTC program on enlistments in ZIP codes, or ZCTAs, and their surrounding areas. We find that there is a positive association between the proximity to a JROTC program and the number of enlistments; however, the larger effects occur in the interactions among the JROTC distance factor, veteran population percentage, and 18- to 24-year-old population size. The combination of higher veteran rates and having a JROTC within a ZIP code is considerably more impactful than the individual factor benefits alone. Additionally, we predict the net improvement associated with establishing a new JROTC in ZIP codes to identify the ZIP codes that are best in terms of increasing the number of enlistments.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

DOD	Department of Defense
DoDI	Department of Defense Instruction
GPA	Grade Point Average
JROTC	Junior Reserve Officers' Training Corps
NCES	National Center of Education Statistics
NDA	National Defense Act
RMSE	Root Mean Square Error
ROTC	Reserve Officers' Training Corps
USPSIG	United States Postal Service Office of Inspector General
ZIP	Zone Improvement Plan
ZCTA	U.S. Census Bureau ZIP Code Tabulation Areas

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The Junior Reserve Officers' Training Corps (JROTC) has played an important role in our society for over a hundred years. The JROTC program has evolved from its origins as a supplemental source of Army Reserve officers to its current mission to improve the character of our nation's youth. The Department of Defense evaluates the success of the program primarily on how it improves the citizenship of its cadets and the schools it supports (Kamarack 2020). However, the JROTC program may provide additional benefit to the military in its ability to help generate enlistments from the community it serves. This study's purpose is to investigate any potential impacts or associations that a JROTC program has on the number of enlistments from that community.

In this study, we utilized multiple regression techniques using mostly publicly available data from 2015 to 2017 containing economic, demographic, education numbers, distances from the nearest military installation and JROTC program, and veteran factors to predict the number of Army enlistments by ZIP code. We limited our analysis to enlistments into the Active Army and the Army Reserves for people aged between 18 and 24. Individual data on cadets or school populations was not available, so the analysis was aggregated to communities at the ZIP code level. Once we developed a model to predict the number of enlistments, we developed an algorithm to predict which ZIP codes could provide the most benefit in terms of increasing the number of enlistments if a JROTC were to be stationed there.

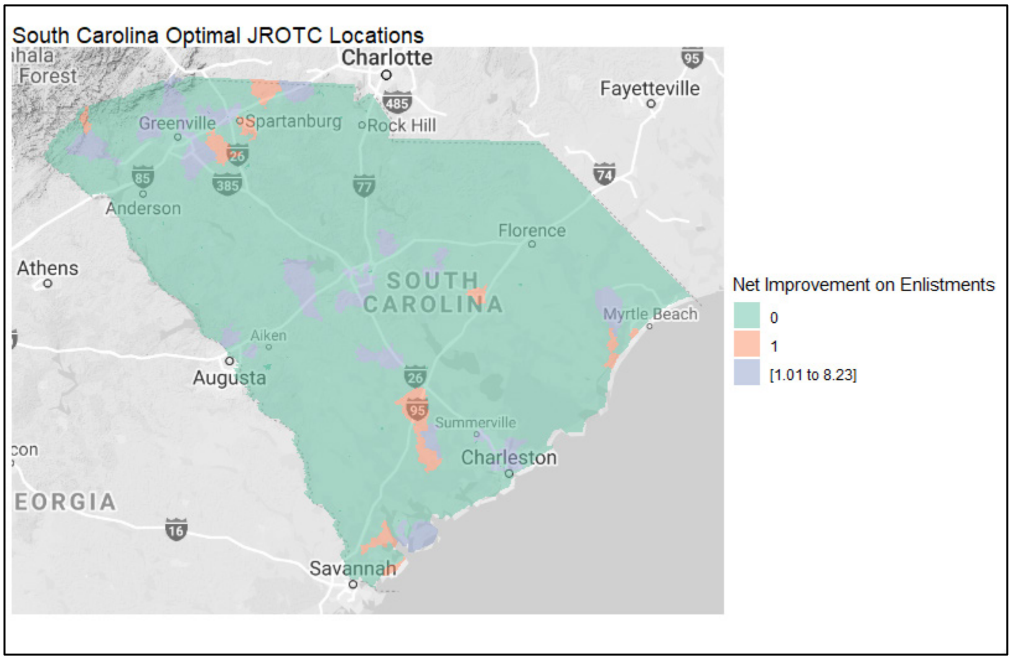
We utilized multiple multivariate linear, Poisson regression, and Negative Binomial regression techniques to identify two fitted models that provided the best predictive power on our 2017 test set. The first model was a multivariate linear regression model with a total of 53 terms containing an intercept, 20 primary effects and 32 two and three-way interactions. The second model was a less complex Poisson regression model with 29 terms containing an intercept, 20 primary effects, and 9 two-way interactions.

The multivariate linear model provided numerous insights into predicting the number of enlistments. Our model determined that there was indeed a positive association between recruiting in a ZIP code and distance to a JROTC program. In other words, ZIP codes with a JROTC program, or close to a JROTC, were predicted to have more enlistments, all other things being equal, while ZIP codes that were further away produce fewer enlistments on average. However, the more important finding was in the interactions between the distance from a JROTC program, veteran population, and 18- to 24-Year-Old population size. We found that in large population areas that were close to a JROTC program, larger veteran populations were associated with a dramatically increased number of enlistments in a ZIP code, compared to their primary effect alone. Another important insight is that the combination of larger veteran populations with higher unemployment rates resulted in a large reduction in the number of enlistments.

The Poisson regression model provide additional insights regarding the effects that varying household income levels had on predicting the number of enlistments. The household income level that showed the largest association to the number of enlistments was the household income between $50,000 and $100,000, which we defined as middle-class. We saw that areas that had a small number of middle-class households produced far fewer enlistments than comparable levels of other income brackets. As the percentage of middle-class populations increased the rate of enlistments increased significantly faster than a linear relationship.

Lastly, we developed an algorithm that predicted the net improvement for each ZIP code utilizing the multivariate linear model. This net improvement was based on calculating the predicted improvement in each individual ZIP code, and surrounding ZIP codes, through the establishment of a new JROTC in that location. Once compiled, we could then visualize this improvement on a state map to highlight areas of a state of interest. An example is shown in Figure 1.

Figure 1. Map of South Carolina illustrating the net improvement in enlistments establishing a JROTC



References

Kamarack K and Arriagea X (2020) Defense Primer: Junior Reserve Officer Training Corps (CRS Report No. IF11313). Accessed from Congressional Research Service website, https://fas.org/sgp/crs/natsec/IF11313.pdf

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PROBLEM DESCRIPTION

The Junior Reserve Officers' Training Corps (JROTC) has played an important role in our society for over 100 years. Congress currently spends over $387 million on this program across the Department of Defense (DOD) enterprise, reaching over 500,000 cadets annually (Kamarck 2020). JROTC's primary purpose is to improve the citizenship of its cadets, instilling in them leadership, confidence, personnel responsibility, and a sense of accomplishment (Army Junior ROTC 2020). An important consideration senior leaders and governmental leadership have is ensuring that the taxpayer money they spend provides significant return on investment. Given that its mission statement is to improve citizenship, the JROTC program has nebulous criteria that make it difficult to quantify its return on investment.

Some studies have shown that cadets do receive numerous benefits like increased grade point average (GPA), reduced juvenile delinquency, and reduced drop-out rate (U.S. Army Junior ROTC 2020). One of the additional benefits that cadets receive when they complete the program is that they receive a higher starting rank if they do choose to join the military. Intuitively, this may play a role motivating cadets to join the military. This additional benefit to the DOD can provide an additional metric to utilize when quantifying the JROTC program's return on investment.

B. PURPOSE

While the primary mission of the JROTC program is to make the youth of America into better citizens, the purpose of this thesis is to investigate the ancillary impact that the JROTC program has on enlistments into the Army (both Regular Army and Army Reserve) using publicly available data. A model that predicts enlistments associated with JROTC can aid decision-makers in quantifying additional impacts of the JROTC program. In addition, it can help narrow the search for potential locations in which to open new JROTC programs, not only assisting in the Army's recruiting mission but also potentially improving the character and citizenship of an area's youth.

C. RESEARCH QUESTIONS:

We seek to determine the impact that a JROTC program has on the number of Army enlistments at the ZIP code level. To that end, we focus on the following questions:

1. Which factors are important when estimating the number of enlistments by ZIP code?
2. Which regression technique provides the most robust model with which to predict the number of enlistments for a ZIP code?
3. What ZIP codes would be enhanced most, in terms of the number of enlistments, through the establishment of a JROTC program?

This study primarily focuses on using regression techniques utilizing data like economic status, demographic information, college enrollment, and proximity to both military installations and high schools with JROTC programs to predict Army recruitment by ZIP Code. All datasets are from the years 2015–2017 and, except for the number of enlistments by age and ZIP code, along with JROTC locations, are publicly available. Only U.S. ZIP codes within the fifty states were considered.

D. SCOPE

1. Assumptions

1. U.S. Census Bureau American Community Survey data estimates portray an accurate estimate of the target actual population.
2. The number of enlistments in each ZIP code is independent of the number of enlistments in other ZIP codes. This assumption may not necessarily be exactly true, as there presumably will be a bleed-over effect between ZIP codes. A few examples include students attending high schools in nearby ZIP codes or that recruiting efforts in one area may spread based on word-of-mouth within a community over areas outside of the local ZIP code.

3. The presence of a JROTC unit will predominately impact the enlistment rate of adults between 18 and 24 as these ages encompass the years immediately following high school graduation.

4. U.S. Census Bureau ZIP Code Tabulation Areas (ZCTA) provide a reasonable analog to their respective ZIP codes to use the demographic data in the analysis.

2. Limitations

1. Lack of available data on individual cadets. The JROTC Headquarters does not maintain individual records on each cadet in the program, partly because of administrative costs, but also to protect the privacy of minors. This limits the ability to study the effectiveness of the JROTC program on enlistments at the individual level. The Army Recruiting Command does have data available on the number of enlistments with JROTC experience, as they typically enlist at higher beginning rank. However, this information is insufficient, as it only provides data on individuals that have had sufficient time in a JROTC program to receive an enhanced rank. Additionally, it does not measure the percentage of JROTC cadets who enlist and do not receive the enhanced rank. This information would provide data only on the direct impact of JROTC on cadets and would ignore the indirect effects on recruitment from other students at the school and in the community as a whole.

2. Data used herein is limited to the number of enlistments for the Regular Army and Army Reserves, as opposed to data on enlistments to all services and components. Additional data on the number of enlistments by age and ZIP code for all the services as well as the Army National Guard would provide a more holistic analysis of the impact on enlistment for the Department of Defense.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND

A. OVERVIEW OF THE JUNIOR RESERVE OFFICERS' TRAINING CORPS

1. Establishment of ROTC and JROTC Programs

The National Defense Act of 1916 established the Senior Reserve Officers' Training Corps, otherwise known as ROTC, and Junior Reserve Officers' Training Corps (JROTC) programs, which operated at the collegiate and high school levels, and provided the Army an officer commissioning source. The ROTC program focused on college-level institutions, where upon graduation, the cadets would commission as a second lieutenant. The JROTC's focus was on secondary schools, where cadets would receive a reserve commission at 21 after completing the program (Coumbe and Harford 1996). Coumbe and Harford (1996) further noted that following World War I, the requirement for many reserve officers was reduced and therefore the commissioning benefit for JROTC cadets was removed. The NDA of 1916 authorized federal military equipment to be loaned to high schools and allowed for active and retired service members to serve as instructors as long as the school followed a prescribed curriculum and maintained an enrollment of over 100 students in the program (Coumbe and Harford 1996).

The JROTC program has faced numerous challenges throughout its history from 1916 until the present day. Only a few years after its founding, World War I began effectively removing most of the financial resources available to fund and operate the program. Financial resources permitted the founding of only 30 units within the first three years of the program (Coumbe and Harford 1996). This inadequate funding limited the expansion of the JROTC program to only 295 units by 1939, over twenty years later. Coumb and Harford (1996) explained that this trend continued throughout the next 50 years, up to the 1990s, tremendously impacting the program's effectiveness and growth. Following World War II the Army "froze JROTC growth due to funding and manpower constraints" (Coumbe and Harford, p. 259).

2. Expansion of the JROTC program in the 1960s

In 1961, Secretary of Defense Robert McNamara attempted to gut the JROTC program and high school units, planning instead to convert them to a competing program, the National Defense Cadet Corps. This began a public outcry to Congress to keep the program, as the public believed the program to be well worth the investment for reducing juvenile delinquency, amongst other things (Coumbe and Harford 1996). This caused Congress to investigate the impact of JROTC and resulted in a published report, finding that while the JROTC program did not serve a direct military purpose as initially intended (commissioning officers), it had many other benefits to both the military and society as a whole. Coumbe and Harford (1996) explain that this caused both a proposed expansion of the JROTC program from 254 units to 2,000 units and the inclusion of the other services, namely the Navy, Marines, and Air Force, sponsoring JROTC programs.

In October of 1963, President Kennedy signed into law the ROTC Vitalization Act of 1964. This act required the Army, Navy, and Air Force to increase the number of JROTC programs they serviced while distributing the JROTC across the country (Coumbe and Harford 1996). The Vitalization act also provided incentives to high schools to hire retired military service as instructors, stipulating that the services would cover half of the personnel costs. Coumbe and Harford (1996) described that the retired service members serving as JROTC instructors would earn a compensation package equivalent to their active duty pay, with all pay and allowances. All other costs would be shared evenly between the school and the respective sponsoring service.

In 1965 the Department of Defense (DOD) strengthened the JROTC program through three initiatives. First, it authorized JROTC cadets to receive either the college level ROTC program or to enlist into the Armed Forces with enhanced rank. Second, the DOD created two distinct tracks within the JROTC program, one focusing on college preparation and the other on military vocational instruction. Third, the DOD mandated that JROTC programs at all non-military high schools be staffed by retired military service members (Coumbe and Harford 1996). These efforts were further expanded in 1970, allowing JROTC cadets to enlist in the Regular Army at the ranks of E2 to E4 depending on experience and performance in the JROTC program.

To further expand the program's student base as well as improve the diversity, the first female cadets entered the program in 1973, in an effort to promote equality in schools. According to a Rand study from 2017, nearly 40 percent of cadets across all services are female (Goldman 2017). While Goldman and his team from Rand acknowledge that this proportion does slightly underrepresent the proportion of females in the population, it is interesting that only 15 percent of the Active Duty force is female. Despite many female cadets not joining the military following graduation, they presumably still learn the positive traits and skills taught within the JROTC program.

3. Expansion and Refocus of JROTC through the 1980s and 1990s

Congress passed Public Law 96-342 in 1980. This reduced the JROTC unit enrollment requirement from having at least 100 participants to having either 100 students, or at least 10 percent of the school's total enrollment, involved with the program (Coumbe and Harford 1996). This effectively allowed more schools to start JROTC programs, as it expanded the ability for schools with smaller populations to qualify. Coumbe and Harford (1996) explained that in the mid-1980s, the JROTC implemented the JROTC Improvement Plan to streamline instruction and management. This plan changed the curriculum to focus less on military training, and instead required 50 percent of the curriculum to focus on technology. This served two functions; first, to increase the number of cadets qualified for science and engineer specialties, for which the Army's Officer Corps had a significant need, and second, to attract cadets who were more academically successful (Coumbe and Harford 1996). In 1987, the JROTC program adopted a new mission statement: "To Motivate Young People to Be Good [later changed to Better] Americans" (Coumbe and Harford, p. 269).

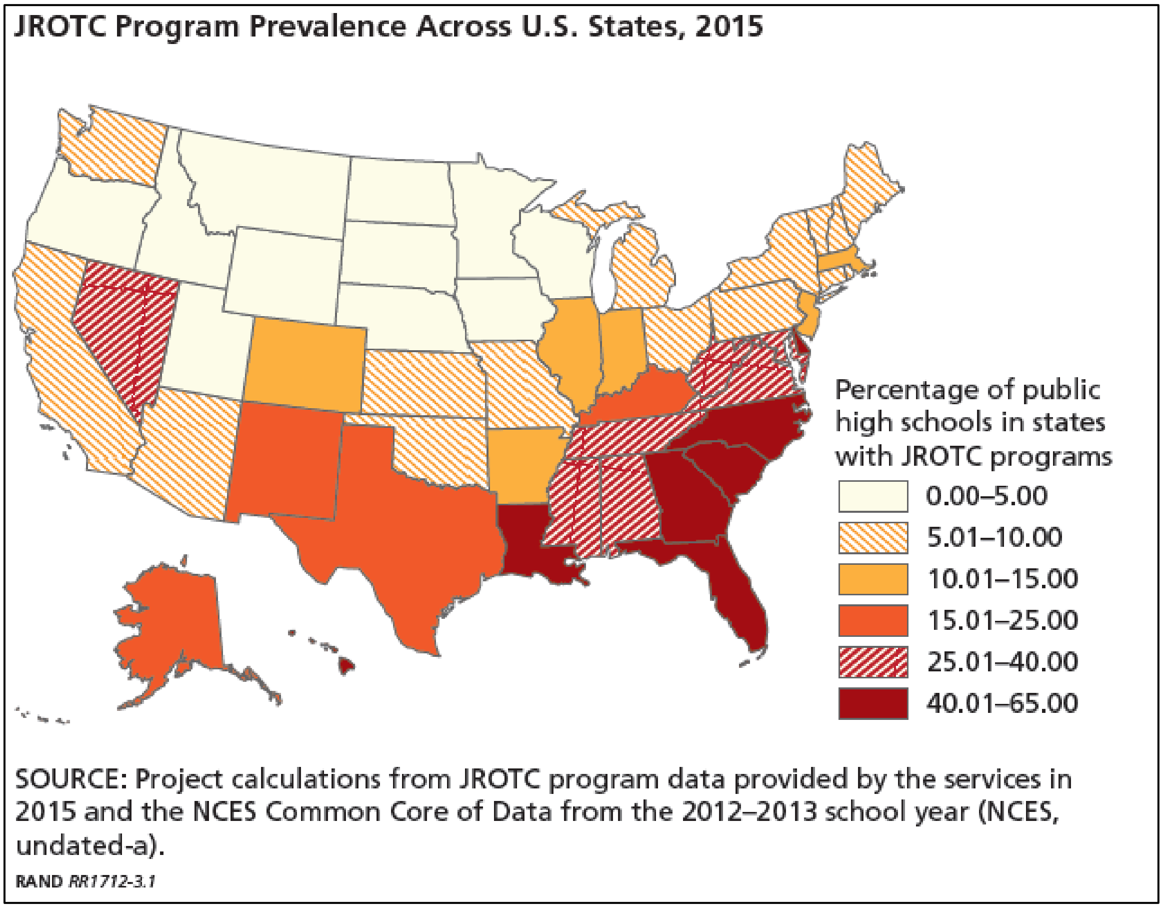
The U.S. Army Cadet Command, which is the higher command under which the JROTC and ROTC programs eventually fell, refocused the JROTC program and instruction under Major General Wallace Arnold's leadership from the late 1980s and on. Coumbe and Harford (1996) described how the JROTC program became a more citizenship-focused program instead of one focused on military training. Throughout the previous twenty years (1970-1990) the JROTC program had grown only moderately.

However, Coumbe and Harford (1996) reported that in 1992 there was a large expansion in the program whereby the JROTC program increased significantly, expanding the number of units by over 60 percent between 1992 and 1996. The public as a whole viewed the JROTC program as an additional benefit the military provided outside of national defense. In August 1992, President George H.W. Bush announced the goal to nearly doubling of the JROTC program from 1,500 to 2,900 units as it was, in his words, a "great program that boosts high school completion rates, reduces drug use, raises self-esteem, and gets these kids firmly on the right track" (Coumbe and Harford 1996, p. 276).

In conjunction with this large growth in the JROTC program, there were two other lines of effort. The first was to provide incentives to disadvantaged schools to open a JROTC program, for which many inner-city schools qualified. The result during this expansion was 35 percent of schools joining the program being able to receive up to 5 years' worth of financial assistance. The second was to expand the geographic diversity in the program (Coumbe and Harford 1996).

Coumbe and Harford (1996) described that the Southeast region of the United States was greatly overrepresented, while other sections of the country were underrepresented. According to a Rand study (2017), the JROTC program to this day serves economically disadvantaged schools at a higher proportion. Further, all four sponsoring services consider the school's Title I status within the criteria for selecting schools to open a new JROTC (Goldman 2017). A school is determined to be a Title I school when the school receives additional funding to support low-income families (National Center for Education Statistics 2020). However, geographic representation is still a significant deficiency for the JROTC program, with the Southeast maintaining a disproportionately higher representation of high schools with a JROTC Program than other regions. Figure 1 from the Rand report shows the geographic overrepresentation of the Southeast.

Figure 1. Percentage of Schools with JROTC Programs. Source: Goldman (2017).



As the JROTC program was expanding, two important efforts were made to bring it into parity with the senior ROTC program. First was the introduction of the JROTC Instructor Course, which helped standardize instruction given at the individual high school level (Coumbe and Harford 1996). Coumbe and Harford (1996) contend that the second was Congress's addition to Title 10 of the United States Code that enabled the federal government to subsidize the annual JROTC summer camps. These summer camps, which had always been an important part of the program, were designed to teach cadets important aspects of leadership and military concepts. Coumbe and Harford (1996) describe how Cadet Command created standardized instruction that would focus on leadership and citizenship, further focusing on improving a cadet's character. Ultimately, Cadet Command's institutional changes and expansion through the 1980s and 1990s evolved

JROTC from its initial goal to provide an avenue to enlistment into the military to a program focused on improving the citizenship of its cadets.

4. Modern JROTC Program

In the Department of Defense Instruction (DoDI) on JROTC in 2006, the DOD declared the intent for the program to "instill in students of American secondary education institutions the following: the value of citizenship, service to the United States, personal responsibility, and a sense of accomplishment." Additionally, the DoDI requires the services to "provide for a fair and equitable distribution of units throughout the nation" (DOD 2006). Each service creates its own criteria when determining where to open the next JROTC. The Army utilizes an Order of Merit system based on numerous criteria, including the school's Title 1 status, student enrollment, and financial circumstances as well as the surrounding area's unemployment rate and illiteracy. The overall distribution of Army JROTCs programs is also factored into the Order of Merit (Kamarck 2020).

The focus of the JROTC program has evolved through its history; however, funding has been the biggest challenge to accomplish its mission. As of 2019, the DOD provided nearly $387 million in total per year to all services for the JROTC program, with over $207 million allocated to the Army program. The estimated enrollment numbers in the JROTC is "more than 500,000 students" (Kamarck 2020). This distills down to the DOD investing nearly $775 per cadet per annum. According to an Information Paper by Cadet Command, approximately 46 percent of Regular Army enlistments come from schools with a JROTC (Dorminey 2019a). Of the approximately 57,500 enlistments per year in the Army, over 26,000 enlistments come from a school with a JROTC. The $387 million per year spent on JROTC's across all services is substantially smaller than the Army Recruiting Command's recruiting and advertising budget of nearly $700 million in 2019 (Myers 2018). Fulton (2016) found that in 2016 the Army spent on average of $22,300 for each recruit enlisted. JROTC may provide a more focused allocation of resources to influence enlistments that may reduce the cost per enlistment.

The Army JROTC program claims multiple benefits to its cadets compared to a school's other students not in the program. These include attendance rates improved by 3 percent, graduation rates nearly 11 percent higher, discipline rates reduced by 3.5 percent, a less than 1 percent drop-out rate compared to overall high school student's drop-out rate of 8 percent, and a 7 percent increase in overall GPA (Army Junior ROTC 2020). This positive effect on JROTC cadets is further supported through a study by the National Association of State Boards of Education (2010). This study found similar improvements in attendance rates, graduation rates, GPA, and also increased SAT and ACT scores.

5. Summary

The JROTC program has changed tremendously since its beginnings in the early 1900s. The program evolved from its initial goal to facilitate entry into the Armed Forces, to improving the character of their cadets, thereby providing benefits to their communities. The continued support and expansion of the programs provide numerous benefits to society that have been recognized by governmental leadership and the public as a whole.

B. LITERATURE REVIEW

Numerous studies have been conducted on the positive impact JROTC programs have on cadets' citizenship in the schools that they operate. This study, however, focuses on the hypothesis that JROTC programs will provide some positive benefit or "lift" to recruiting. To determine how beneficial JROTC is in providing lift to recruiting, the following studies provide some context to important factors in recruiting:

In his master's thesis Fulton (2016) conducted an analysis using R, a statistical language and environment, to predict leads from publicly available data resources to reduce the dependence on commercial services for the number of leads from a ZIP code (R Core Team 2019). Fulton clustered ZIP codes based on demographic, educational, health-related, military factors, and higher education factors containing over 347 variables. He utilized the R package *treeClust* to group the ZIP codes into groups that performed better than some commercial services to predict the number of leads generated for each ZIP code (Buttrey 2018). His efforts yielded that the cluster assignments based on economic activity provided the highest predictive power utilizing Poisson regression (Fulton). However, he did not include the presence of a JROTC in his analysis.

In an Institute for Defense Analyses paper on Geographic Diversity in Military Recruiting, the project team utilized machine learning analysis on total accessions for males, females, and quality levels (Goldberg 2018). They conducted their analysis at the county level utilizing demographic, economic, local postsecondary educational opportunities, JROTC locations, health, veteran populations, and Armed Services Vocation Aptitude Battery (ASVAB) Career Exploration Program data. Through their analysis they found that size of the veteran population had a strong effect on the number of accessions, while the percentage of people attending college had a strong negative effect. Their estimates also found that the presence of JROTC programs within the county had a moderately positive impact. Goldberg and his team (2018) also conducted a regional trend analysis on the data that highlighted differences among geographic regions of the country.

Dorminey (2019a) from Cadet Command studied the impact of JROTC programs on Army enlistments. She utilized DOD service data that provided the JROTC program locations, accession data based on enlistments, schools and education history, and data from National Center for Education Statistics (NCES) on enrollment and school demographics to compare mean enlistment rates based on presence or absence of JROTC by sponsoring service. She conducted a comparison of mean enlistment rates for Regular Army, Army Reserve, and Army National Guard. The results indicated that students at Army JROTC programs were twice as likely to enlist in the Army than those in a school with no JROTC program. Additionally, non-Army JROTC programs generated Army enlistments in numbers 60–80 percent larger than schools with no JROTC. Dorminey (2019a), however, did acknowledge that the analysis did not explore if JROTC presence caused the increase of enlistments or if the presence was simply correlated with higher enlistments.

III. DATA DESCRIPTION

This chapter describes the various data sources utilized to build the dataset to model the number of enlistments for a ZIP code for years 2015–2017. All data was from publicly available sources on government agency websites, except for the actual number of enlistments in both the Regular Army and Army Reserves by age and ZIP code and locations of the JROTC units. The U.S. Army Recruiting Command provided the number of enlistments by age and zip code in the form of aggregate numbers. Secondary school locations with senior populations greater than twenty, and JROTC status by service, was provided by the U.S. Army Cadet Command and came from the same dataset that is utilized in its comparison of means analysis described in the Literature Review section.

Cadet Command utilized a peer-reviewed methodology, obtaining its data from the National Center for Education and Statistics (NCES), a public governmental agency website. The JROTC program locations were provided by the Army, Navy, Air Force, and Marine, and U.S. Army Recruiting Command provided the number of enlistments recruited from each of the 19,101 high schools. Recruiting Command also included the latitude and longitude of the nearest military installation. All other datasets utilized were collected from governmental agencies including the United States Census Bureau and NCES.

A. ANALYSIS AT ZIP CODE LEVEL

Zone Improvement Plan (ZIP) codes were initially introduced by the United States Postal Service in 1963 (United States Postal Service Office of Inspector General [USPSIG] 2013). The code's purpose was to aid the USPS to automate its sorting and delivery methods. In USPSIG's report (2013), it contended that an accidental benefit of "digitizing" the geographic areas of the country ZIP codes is a numerical representation of areas of the country. Many industries adopted ZIP codes to organize their data according to these geographical boundaries. Examples of some industries that utilize ZIP codes in their operations today include local and federal agencies, utilities, financial institutions, real estate, marketing, and universities (USPSIG 2013). The most important agency in terms of data gathering for this study is the United States Census Bureau. The Census Bureau utilizes ZIP codes as the lowest levels of aggregation for demographic information. There are currently over 41,000 ZIP codes; however, many are merely administrative, dedicated to individual businesses, or represent areas with zero population.

The U.S. Census Bureau, in order to generalize these USPS ZIP code areas (which are mail delivery routes), to areal features, established ZIP code Tabulation areas (ZCTA). Essentially, the ZCTAs aggregates the residential and non-residential addresses from the ZIP codes into blocks, and then the Bureau aggregates these blocks into the most frequently used ZIP code in those blocks and assign that block to a single ZCTA (United States Census Bureau 2015).

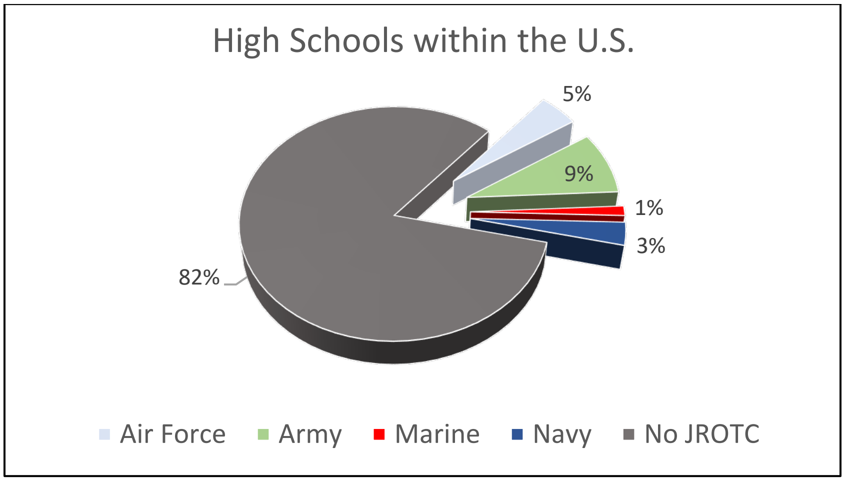
B. JROTC UNITS

In 2017, there were a total of 3,348 active JROTC units in both public and private high schools, with each service sponsoring a portion of these JROTC units (Dorminary 2019b). Table 1 provides the distribution of JROTC units for the United States high schools with over 20 seniors and Figure 2 graphically depicts the relative proportion of these schools that have JROTCs. The Army, by far, has the largest proportion of JROTC units, with the Air Force following as the next largest. The Marines and Navy have the smallest number of JROTC units, equaling the number of Air Force units when combined. High schools that have JROTC units account for fewer than 18% of all high schools in the United States.

Table 1. Number of JROTC Units by Service. Source: Dorminary (2019b).

Service	Sponsored Units
Air Force	869
Army	1,631
Marine	251
Navy	598
No JROTC	15,752
Total Schools	19,101

Figure 2. Distribution of JROTC Units in High Schools. Source: Dorminary (2019b).



C. DESCRIPTION OF VARIABLES

1. Response Variable

The response variable for the analysis is the number of enlistments by ZIP code between the ages of 18–24. While the age of enlistees ranged from 16 to 49 in the Regular Army and Army Reserves, we limited the scope under the assumption that the number of enlistments impacted by a JROTC program will be seen in high school seniors and for only a few years after graduation. The number of enlistments for the Regular Army and Army Reserves was combined for this analysis (U.S. Army Recruiting Command 2019).

2. Demographic Data

Demographic data was gathered from the U.S. Census Bureau American Community Survey datasets utilizing the Census Bureau's API via the R package *acs* (Glenn 2019). These datasets were tabulated at the national level by 5-digit ZIP Code Tabulation Areas. This resulted in 33,120 total ZIP codes with demographic data. Approximately 8,000 other ZIP codes had zero population. Based on previous work, we focused the dataset on demographic such as based on race, economic factors, veteran population, veteran economic data, unemployment rate, and 18–24 age population (United States Census Bureau 2019). The NCES data was collected for the total population enrolled in a postsecondary school as a full-time student (National Center for Education Statistics 2019). This was further aggregated to the ZIP code level.

3. Geographic Variables

Two categorical variables were created defining ZIP code from their respective state and further assigning them one of four regions within the country. We utilized the same methodology used in the Rand study to determine region assignment by state. (Goldman 2017). The assignment matrix is illustrated in Table 2.

Table 2. Assignment of States to Regions

Region Designation							
New England		**South**		**Midwest**		**West**	
CT	RI	AR	PR	IA	ND	AK	NM
MA	VT	FL	SC	IL	NE	AZ	NV
ME		GA	TX	IN	OH	CA	OR
NH		LA	VA	KS	SD	CO	UT
NJ		MS	WV	MI	WI	HI	WA
NY		NC		MN		ID	WY
PA		OK		MO		MT	

4. JROTC and Military Categorical Variables

To form Military and JROTC categorical variables, we used the Cadet Command dataset to find the ZIP code for each high school, employing the function "revgeocode" from the R package *ggmap* to find the address for each school based on its latitude and longitude (Kahle and Wickham 2020). Using the R package *ZipRadius*, which returns a list of all ZIP codes within a specified mile radius, we assigned factor values for each ZIP code depending on its distance from the nearest ZIP code with a JROTC for all 33,120 ZIP codes from the Census Bureau dataset (Ewing 2020). The schema used to assign values is depicted in Table 3; the lowest applicable value takes precedence for each ZIP code.

Table 3. Defining JROTC Factor Levels

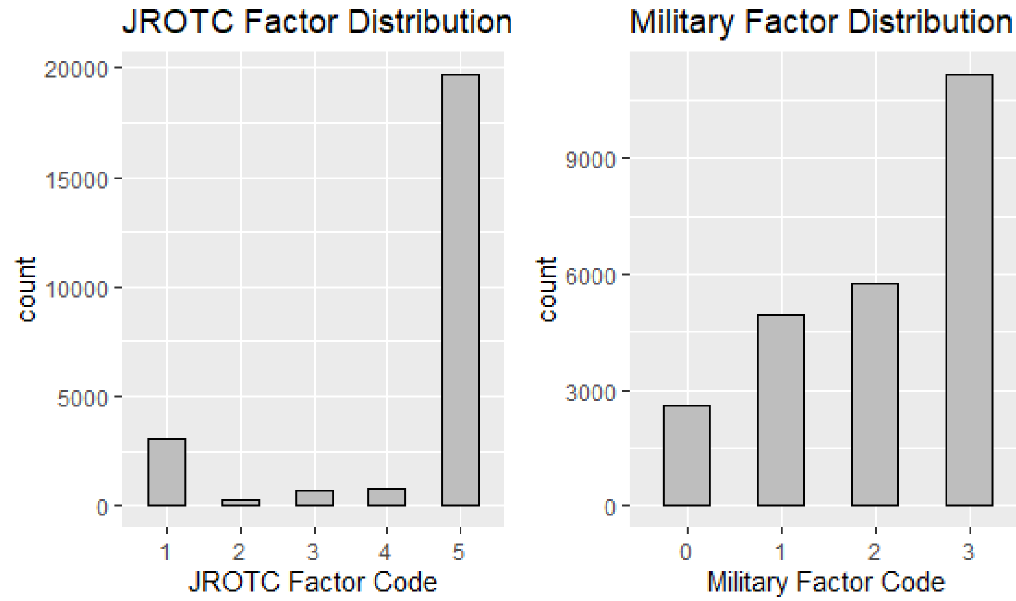
JROTC Factor	
Value	Description
1	ZIP code contains JROTC
2	JROTC within 5 miles of ZIP code
3	JROTC between 5 and 10 miles of ZIP code
4	JROTC within 10 and 15 miles of ZIP code
5	No JROTC within 15 miles

Military Factor codes were assigned similarly. To identify the nearest military installations, we utilized Opendatasoft to download data for over 450 active duty military installations; this data contained the latitude and longitude of the base, shapefile coordinate system, state, service, and other relevant information (Opendatasoft 2020). Each installation was found via reverse geocode lookup and values for each nearby ZIP code assigned based on the results of that lookup. The schema used to assign values for Military Factors is depicted in Table 4. Figure 3 depicts the distribution of the number of ZIP codes for both JROTC and Military Factor codes. We can see the vast majority of ZIP codes do not have a JROTC program nearby; however, a significant proportion of ZIP codes do have a JROTC program in the ZIP code. We can also see that the majority of ZIP codes do not have a Military base near them, but these codes are more evenly distributed across the other factor levels than the JROTC ones.

Table 4. Defining Military Factor Levels

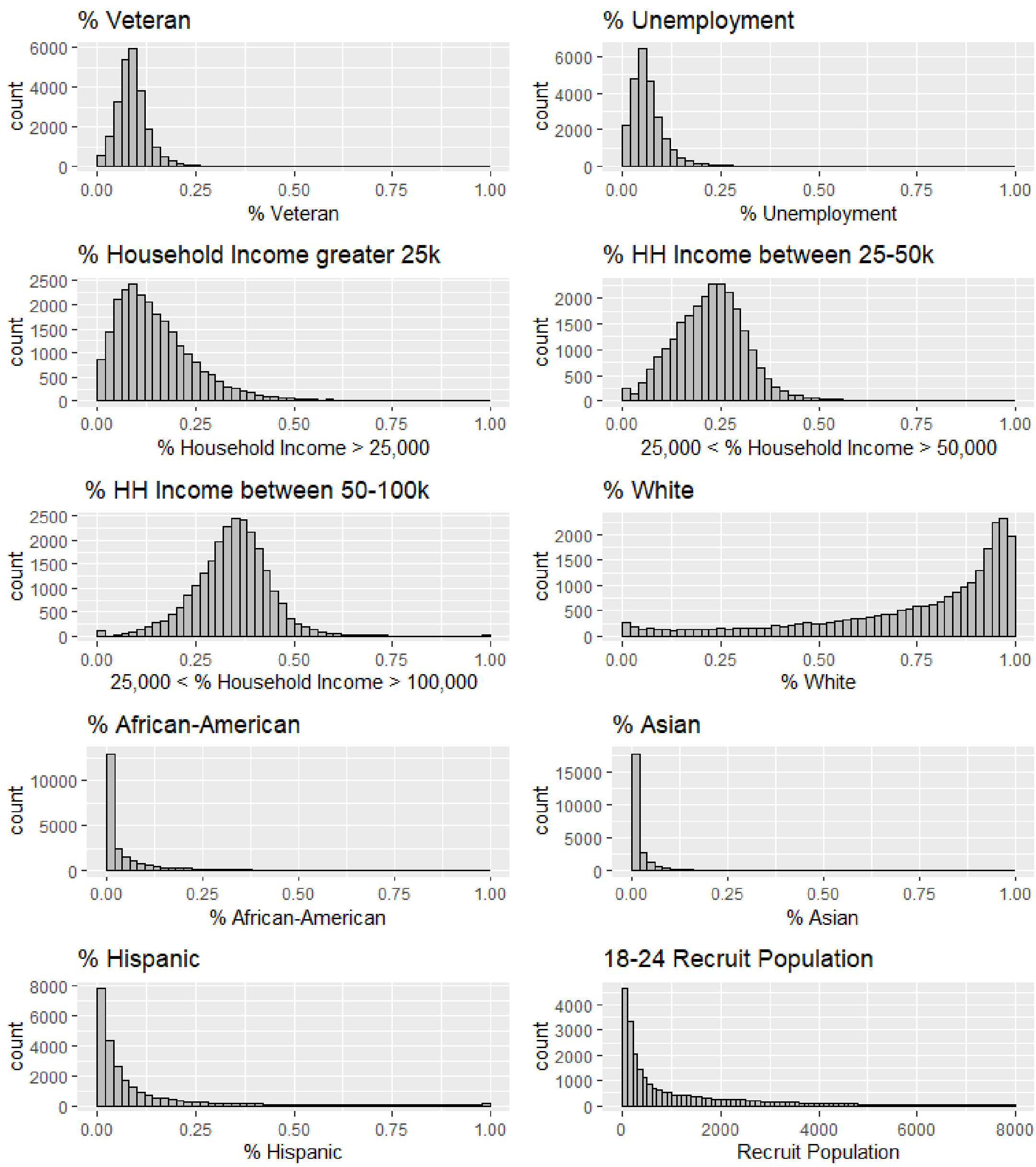
Military Factor	
Value	Description
0	Military Base within 10 miles of ZIP code
1	Military Base between 10–25 miles
2	Military Base between 25–50 miles
3	Military Base > 50 miles away

Figure 3. Distribution of JROTC and Military Factors



To provide some context for the values for the predictor variables, Figure 4 depicts the distributions for some of the demographic variables that will be used in our analysis.

Figure 4. Examples of Demographic Variables Distributions. Source: United States Census Bureau (2020).



THIS PAGE INTENTIONALLY LEFT BLANK

IV. METHODOLOGY

A. INTRODUCTION

We investigated three different regression methods to determine which regression modeling method provides the best predictive capability for estimating the number of enlistments generated from each ZIP code for the Regular Army and Army Reserve. We split the dataset into a training set, which contained the data from 2015 and 2016, and a test set containing the 2017 data.

B. MULTIVARIATE LINEAR REGRESSION

We utilized stepwise AIC method for variable selection to narrow our focus on important variables to use in building the Linear Regression model. We used an initial full factorial design with three-way interactions containing 91 primary effects and thousands of two-way/three-way interactions. Ultimately, we chose a model with 21 main effects and only a small subset of the two-way and three-way effects to balance the effectiveness and the complexity of the model. Table 5 depicts the primary effects chosen with their respective variable type.

Table 5. Defining Predictor Variable Types

Primary Effect Variable Selected	Variable Type
JROTC factor	Categorical
Military Factor	Categorical
Region	Categorical
Percent Veteran Population	Numeric
Percent White	Numeric
Percent African American	Numeric
Percent Asian	Numeric
Percent Total Hispanic	Numeric
Percent Household Income < 25,000	Numeric
Percent 25,000 < Household Income < 50,000	Numeric
Percent 50,000 < Household Income < 100,000	Numeric
Percent Unemployment	Numeric
# of Full-time Student	Numeric
18-24 Recruit Population	Numeric

Table 6 depicts the full model chosen with any necessary transformations applied and respective coefficients. The summary of the results shows an adjusted R squared of .649. We used the square root of the number of enlistments for a ZIP code as the response variable, as the transformation significantly improves model diagnostics and reduce the spread of the count data. The resulting model contains a total of 52 predictor variables made up of 21 main effect variables including all categorical variables, 21 two-way interactions, and 10 three-way interactions. The model performed very well in terms of diagnostics. The base factors in the intercept are JROTC Factor 1 representing that the ZIP code has a JROTC, Military Factor Code 0 representing that there is a Military installation within 10 miles of the ZIP code, and the Region is the Midwest.

Table 6. Multivariate Linear Regression Model

Linear Regression Model	**Square Root of Number of Recruits**				
Predictors	*Est.*	*CI*			*p*
Intercept	–0.03	–0.40	—	0.33	0.852
JROTC Factor = 2	–0.10	–0.67	—	0.47	0.735
JROTC Factor = 3	–0.87	–1.31	—	–0.42	**0.001**
JROTC Factor = 4	–0.40	–0.86	—	0.06	0.088
JROTC Factor = 5	–0.47	–0.77	—	–0.17	**0.002**
Military Factor Code = 1	0.25	0.04	—	0.47	**0.022**
Military Factor Code = 2	0.36	0.15	—	0.58	**0.001**
Military Factor Code = 3	0.17	–0.03	—	0.37	0.09
Region = NE	–0.01	–0.03	—	0.01	0.471
Region = South	0.12	0.11	—	0.14	**<0.001**
Region = Western	0.06	0.04	—	0.08	**<0.001**
% HH Income < 25k	0.16	0.08	—	0.24	**<0.001**
% 25k < HH Income < 50k	0.15	0.08	—	0.23	**<0.001**
% 50k < HH Income < 100k	–0.05	–0.30	—	0.20	0.689
% Unemployment	–0.45	–0.89	—	–0.01	**0.047**
% White	0.53	0.34	—	0.72	**<0.001**
% African Amer.	0.28	0.19	—	0.37	**<0.001**
% Asian	0.25	0.08	—	0.41	**0.003**
% Hispanic	0.34	0.25	—	0.43	**<0.001**
Sixth Root of # of Full Time Student	–0.48	–0.52	—	–0.45	**<0.001**
Cube Root of % Veterans	–0.50	–1.29	—	0.29	0.214
Square Root of 18–24 Recruit Population	–0.01	–0.02	—	–0.01	**<0.001**
JROTC Factor = 2 * Cube Root of % Veterans	0.74	–0.51	—	1.99	0.244
JROTC Factor = 3 * Cube Root of % Veterans	2.05	1.07	—	3.02	**<0.001**
JROTC Factor = 4 * Cube Root of % Veterans	0.75	–0.28	—	1.78	0.155
JROTC Factor = 5 * Cube Root of % Veterans	0.82	0.14	—	1.50	**0.018**
% White * Cube Root of % Veterans	–0.11	–0.54	—	0.32	0.622
% 50k < HH Income < 100k * Cube Root of % Veterans	–0.40	–0.96	—	0.16	0.162
% Unemployment * Cube Root of % Veterans	1.36	0.37	—	2.34	**0.007**

Linear Regression Model	Square Root of Number of Recruits				
Military Factor Code = 1 * Cube Root of % Veterans	–0.85	-1.31	—	–0.38	**<0.001**
Military Factor Code = 2 * Cube Root of % Veterans	–1.18	-1.64	—	–0.72	**<0.001**
Military Factor Code = 3 * Cube Root of % Veterans	–0.76	–1.18	—	–0.33	**<0.001**
JROTC Factor = 2 * Square Root of 18–24 Recruit Population	0.01	0.00	—	0.02	0.067
JROTC Factor = 3 * Square Root of 18–24 Recruit Population	0.04	0.03	—	0.05	**<0.001**
JROTC Factor = 4 * Square Root of 18–24 Recruit Population	0.02	0.01	—	0.03	**<0.001**
JROTC Factor = 5 * Square Root of 18–24 Recruit Population	0.03	0.02	—	0.03	**<0.001**
% White * Square Root of 18–24 Recruit Population	–0.04	–0.04	—	–0.03	**<0.001**
% 50k < HH Income < 100k * Square Root of 18–24 Recruit Population	0.02	0.01	—	0.03	**<0.001**
% Unemployment * Square Root of 18–24 Recruit Population	0.13	0.10	—	0.15	**<0.001**
Military Factor Code = 1 * Square Root of 18–24 Recruit Population	–0.02	-0.03	—	–0.02	**<0.001**
Military Factor Code = 2 * Square Root of 18–24 Recruit Population	–0.03	–0.03	—	–0.02	**<0.001**
Military Factor Code = 3 * Square Root of 18–24 Recruit Population	–0.02	–0.02	—	–0.01	**<0.001**
Cube Root of % Veterans * Square Root of 18–24 Recruit Population	0.12	0.11	—	0.14	**<0.001**
JROTC = 2 * Cube Root of % Veterans * Square Root of 18–24 Rec. Pop	–0.04	–0.07	—	–0.02	**<0.001**
JROTC = 3 * Cube Root of % Veterans * Square Root of 18–24 Rec. Pop	–0.10	–0.12	—	–0.08	**<0.001**
JROTC = 4 * Cube Root of % Veterans * Square Root of 18–24 Rec. Pop	–0.04	–0.07	—	–0.02	**<0.001**
JROTC = 5 * Cube Root of % Veterans * Square Root of 18–24 Rec. Pop	–0.07	–0.08	—	–0.06	**<0.001**
% White * Cube Root of % Veterans * Square Root 18–24 Rec. Pop.	0.04	0.02	—	0.05	**<0.001**
% 50k < HH Income < 100k * Cube Root of % Veterans * Square Root of 18–24 Recruit Population	0.06	0.04	—	0.09	**<0.001**
% Unemp. * Cube Root of % Veterans * Square Root 18–24 Rec. Pop.	–0.33	–0.38	—	–0.27	**<0.001**
MFC= 1 * Cube Root of % Veterans * Square Root of 18–24 Rec. Pop	0.07	0.06	—	0.08	**<0.001**
MFC= 2 * Cube Root of % Veterans * Square Root of 18–24 Rec. Pop	0.07	0.06	—	0.09	**<0.001**

Linear Regression Model	Square Root of Number of Recruits			
MFC= 3 * Cube Root of % Veterans * Square Root of 18–24 Rec. Pop.	0.05	0.04 —	0.06	**<0.001**

Table 7. Performance of Linear Model

Performance on Training	
Observations	48,988
R^2 / adjusted R^2	0.649 / 0.649
F(52,48963)	1744.8083
Performance on Test Set	
Observations	24,487
Mean Abs. Error Transformed Response	0.5677
RMSE Transformed Response	0.7000
R^2 on Test Set	0.6924
Mean Abs. Error Untransformed Response	1.5296
RMSE Untransformed Response	2.8366
R^2 on Test Set Untransformed Response	0.6182

C. POISSON REGRESSION MODEL

As the number of enlistments for a ZIP code is a type of count data, utilizing a Poisson regression model seems like a natural fit to predict future enlistments. Poisson regression coefficients typically represent a change in the log of the response variable. This reflects the assumptions in the Poisson model that responses have approximately equal means and variances. One important consideration for Poisson regression is the presence of a large number of zeros within the response variable. In order to utilize Poisson regression or Negative Binomial regression, which differs from the Poisson as it relaxes the requirement that the mean equal the variance, we must assume the number of enlistments from a ZIP code can be described by a Poisson or Negative Binomial distribution, conditional on the predictors.

To investigate the effectiveness of the Poisson regression methodology, we developed two models, each using all primary effects from the same dataset used in the linear model and only a portion of two-way interactions that significantly improved the performance of the models. The first model created is a standard Poisson regression, with a log transformation on the 18- to 24-Year-Old factor. The second model utilized an offset parameter on the log of the 18–24 population of a ZIP code (Rodriguez 2016). This offset is a way to force the coefficient for that factor to be fixed at 1, which works under the assumption that the rate based on population follows the same slope, allowing differences in population sizes across ZIP codes to be constant in terms of their ratio. Table 8 depicts the Poisson model without the offset parameter, and Table 9 shows a comparison of both models' effectiveness in predicting the number of enlistments for the 2017 test set.

Table 8. Poisson Regression Model without offset parameter

Poisson Regression	**Log(Number of Recruits)**					
Predictors	*Log Mean*	*std. Error*	*CI*			*p*
(Intercept)	–5.77	0.07	–5.91	—	–5.64	**<0.001**
JROTC Factor = 2	0.07	0.03	0.01	—	0.13	**0.024**
JROTC Factor = 3	0.16	0.02	0.11	—	0.21	**<0.001**
JROTC Factor = 4	–0.01	0.03	–0.07	—	0.06	0.832
JROTC Factor = 5	0.05	0.01	0.03	—	0.08	**<0.001**
Military Factor Code = 1	–0.20	0.02	–0.23	—	–0.17	**<0.001**
Military Factor Code = 2	–0.16	0.02	–0.21	—	–0.12	**<0.001**
Military Factor Code = 3	–0.21	0.02	–0.25	—	–0.17	**<0.001**
Region = NE	–0.06	0.01	–0.08	—	0.04	**<0.001**
Region = South	0.22	0.01	0.20	—	0.23	**<0.001**
Region = Western	0.15	0.01	0.13	—	0.16	**<0.001**
# of Full Time Student	–0.22	0.01	–0.24	—	–0.21	**<0.001**
% Veterans	5.83	0.10	5.62	—	6.03	**<0.001**
% HH Income < 25k	–0.07	0.05	–0.16	—	0.02	0.110
% 25k < HH Income < 50k	0.46	0.05	0.36	—	0.55	**<0.001**
% 50k < HH Income < 100k	1.78	0.05	1.69	—	1.88	**<0.001**
% Unemployment	0.27	0.09	0.09	—	0.46	**0.003**

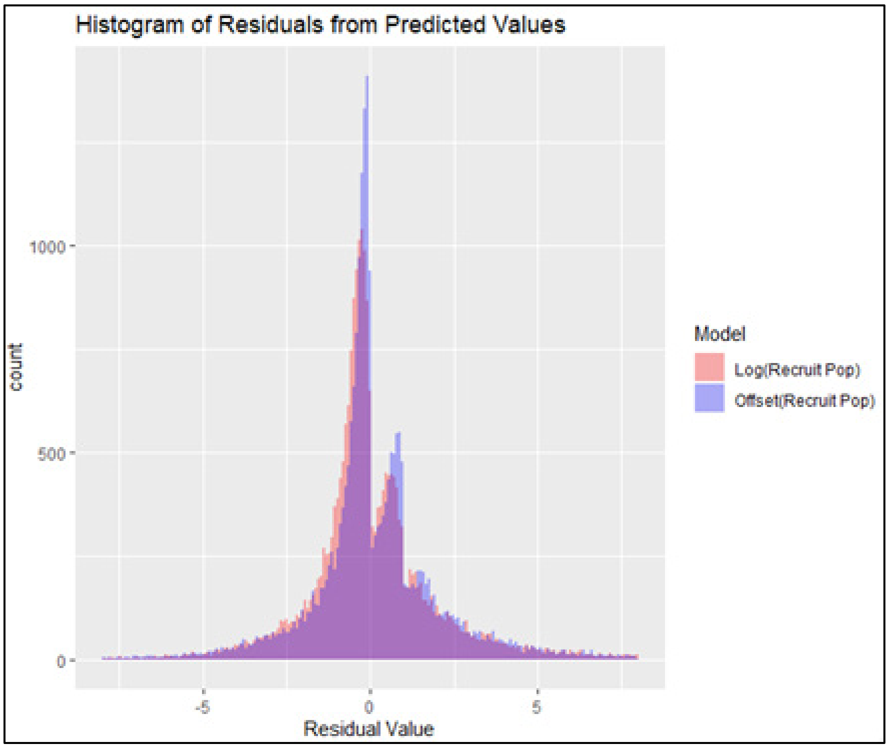
Poisson Regression	**Log(Number of Recruits)**					
% White	–0.15	0.06	–0.26	—	–0.03	**0.011**
% African Amer.	0.45	0.06	0.34	—	0.57	**<0.001**
% Asian	0.19	0.08	0.04	—	0.35	**0.015**
% Hispanic	0.47	0.06	0.35	—	0.58	**<0.001**
Log(18-24 Recruit Population)	0.79	0.01	0.78	—	0.80	**<0.001**
JROTC Factor = 2 * % Veterans	–2.68	0.22	–3.12	—	–2.25	**<0.001**
JROTC Factor = 3 * % Veterans	–3.81	0.20	–4.21	—	–3.42	**<0.001**
JROTC Factor = 4 * % Veterans	–1.71	0.30	–2.30	—	–1.13	**<0.001**
JROTC Factor = 5 * % Veterans	–3.14	0.13	–3.40	—	–2.89	**<0.001**
Military Factor Code = 1 * % Veterans	3.45	0.14	3.17	—	3.73	**<0.001**
Military Factor Code = 2 * % Veterans	2.96	0.22	2.50	—	3.38	**<0.001**
Military Factor Code = 3 * % Veterans	2.93	0.19	2.56	—	3.29	**<0.001**

Table 9. Performance Comparison of Poisson Models

	Poisson Regression	Poisson with offset
Degrees of Freedom	48461/48433	48461/48434
Mean Absolute Error	1.5531	1.5649
RMSE	2.7751	2.9332
LogLik	85158.436	87333.06
McFadden Pseudo R^2	0.4875	0.4744

As seen in Table 9, the inclusion of the offset reduced the effectiveness of the model, with the non-offset model having a RMSE that is smaller. We can see that the recruit population factor had a coefficient of 2.2 in the unrestricted model, whereas in the offset model that coefficient is forced to be 1. This is likely the cause of the difference between the models' predictive power. One potential explanation is that the population size coefficient is not consistent across all population sizes. The U.S. Army Recruiting Command in part prioritizes its distribution of recruiters to areas with larger populations, thereby potentially enhancing the rate of enlistments with larger populations. This would then cause the coefficient for the log of recruit population to change throughout population sizes. Figure 5 plots the residuals between predicted and observed values for 2017 test set. The histograms seem to indicate that the offset model seems to underestimate the number of enlistments compared to the other model.

Figure 5. Comparison of Actual-Predicted Values for Poisson Models



D. ZERO-INFLATED MODELS

We tested a zero-inflated Poisson regression model; this model separates the zero and non-zero cases and models them separately. We used the "zeroinfl" function from the *pscl* package in R to fit the zero-inflated Poisson model (Jackman 2020). The zero-inflated Poisson regression model has two different sets of coefficient estimates: one represents a model with a binomial logit link to predict whether the result would be a zero, and the other representing the log-link Poisson model. An additional benefit to the zero-inflated model is that it provides for overdispersion, relaxing the Poisson requirement that the mean be equal to the variance. Table 10 compares the RMSE and Mean Absolute Error of the

Poisson Regression model against the Zero-Inflated Poisson Regression Model on the 2017 test set. We can see that each model performs better in one of the measures of effectiveness. This indicates that the standard Poisson Regression model performs reasonably well predicting if the number of enlistments from a ZIP code was zero.

Table 10. Performance of Poisson Regression versus Zero-Inflated Poisson Regression on Test Set

	Poisson Regression	Zero-inflated Poisson
Mean Absolute Error	1.5531	1.5648
RMSE	2.7751	2.7918

E. NEGATIVE BINOMIAL REGRESSION MODEL

An alternative to the Poisson Regression is the Negative Binomial Regression, which relaxes the assumption of that the mean be equal to the variance with the inclusion of a dispersion parameter to account for this difference. To determine if a Negative Binomial Model is better for predicting the number of enlistments than the Poisson or Linear Regression Models, we created two models that used the same factors used in the Poisson regression with and without the offset parameter for the population. As in the Poisson regression models, the use of an offset for the population covariate reduced the models' effectiveness. As with the Poisson regression models, the non-offset Negative Binomial model performed better across the board; however, the RMSE is considerably worse with the offset implementation. Ultimately, the Poisson Regression model performed better than the Negative Binomial model, as summarized in Table 11.

F. COMPARING REGRESSION MODELS

Ultimately, we want to compare how effective each of the model types performed on predicting the number of enlistments by ZIP code. The Poisson and Negative Binomial regressions performed better when each did not utilize an offset variable so these two models will be compared against the linear model. Overall, the three models all performed similarly in terms of diagnostics; however, we cannot compare the R squared of the linear model against the Poisson and Negative Binomial as the pseudo R squared values provided does not represent the same thing. Table 11 depicts the effectiveness of each model against the 2017 test set. In order to compare with equivalent units, we square the predicted value for the Linear regression model before comparing against the actual values so all residuals are in terms of the number of enlistments. We can see the Poisson regression performed the best in terms of root mean squared error for all three models. Additionally, the Poisson regression model performed slightly better than the Negative Binomial regression model in absolute mean error. However, the Linear regression model performed the best of all three models in terms of absolute mean error.

The Poisson and Negative Binomial performed very similarly in terms of residuals, which is not surprising as both models are nearly identical in framework except for the assumption of mean equaling variance. However, the Linear model generally seemed to underestimate the number of enlistments as the residuals seemed to have a stronger positive skew, while the other two overestimated the number of enlistments having a stronger negative skew.

Table 11. Comparison of Regression Models on Test Set

	Poisson Regression	Neg. Binom. Regression	Linear Regression
Mean Abs. Error	1.5531	1.5594	1.5296
RMSE	2.7751	2.8450	2.8366
Model Complexity	29 terms	29 terms	53 terms

Another important aspect of comparing these models is model complexity. All the models perform reasonably well, but the Linear regression model is far more complex with many more two and three-way interactions reducing the model's explanatory benefits. An additional component to a model's complexity is how many variable transformations exist or are required. The Linear model required transformations of the predictors for full-time students, recruit population, veteran population, and most importantly the number of enlistments. These complexity issues effectively reduce the ability to clearly understand how each of those predictor variables impacts the response.

V. RESULTS

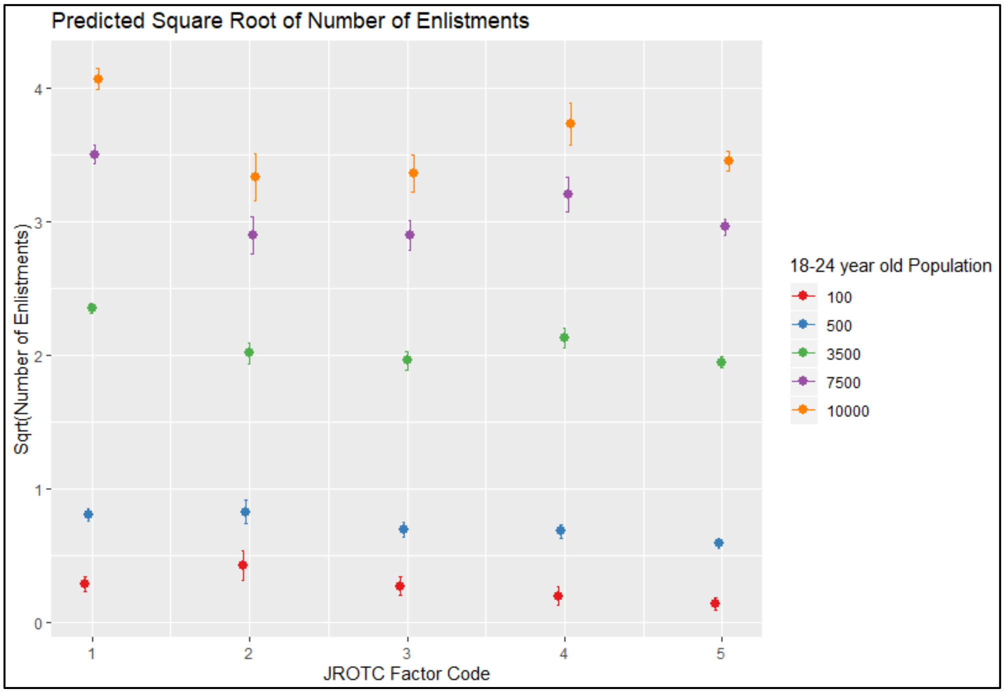
A. MULTIVARIATE LINEAR REGRESSION

The goal of this study was to investigate the effects that JROTC has upon the number of enlistments in a ZIP code. Logically, we can infer that larger populations will likely be associated with larger numbers of enlistments. Therefore, we will present the associations on the number of enlistments in the context of various 18–24 Recruit population sizes.

1. Distance from a JROTC

Figure 6 illustrates the how the presence and distance from a JROTC program has on the square root of Number of Enlistments. Recall from Table 3, which defined the JROTC Factor codes, that a value of 1 corresponds to a ZIP code having a JROTC program and as the value becomes larger the farther away the nearest JROTC program is. There are five colors representing five different 18- to 24-Year-Old population sizes and the figure shows, from left to right, how increased distance from a JROTC program is associated with each population differently in terms of enlistments. We see that in the smaller populations, represented by the red and blue points, the effect associated with an increase in distance from a ZIP code to its nearest JROTC behaves differently than with larger population sizes. In the smaller population sizes, we see a slight positive effect when the JROTC program is within 5 miles of the ZIP code. This may illustrate small-town America, where students are bused further away from their homes to high schools. However, as the population size grows, the further away a community is from a JROTC, the number of enlistments drops significantly indicating the JROTC program's association in dense populations are more localized to that community. An interesting observation is that in the larger population sizes there is a greater rate of enlistment for locations having a JROTC within 10–15 miles than for those where a JROTC is located at the 5–10 miles level, which may indicate other factors like economic data becoming more influential on the number of enlistments.

Figure 6. Effect of Distance from a JROTC on Enlistments on Different Population Sizes

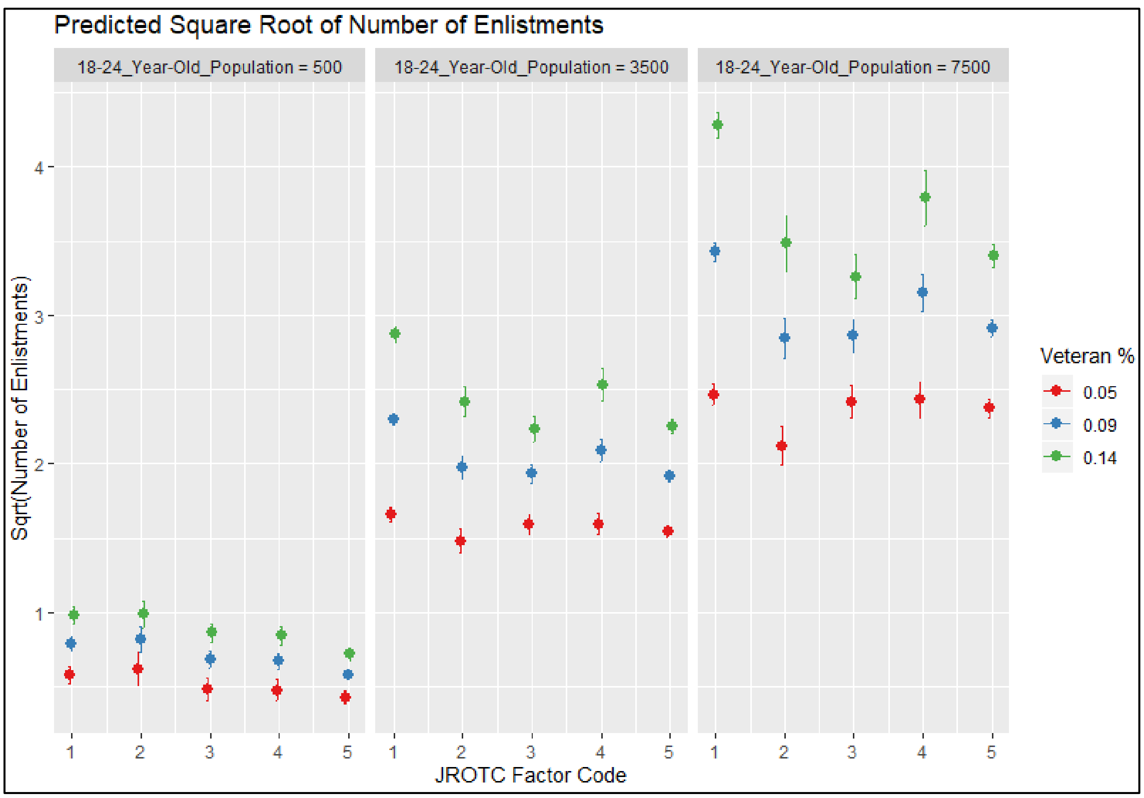


2. Interaction between Veteran Population and distance from a JROTC

The linear model indicated that two-way interactions between JROTC Factor Code and Veteran population are influential in predicting the number of enlistments. Figure 8 depicts that relationship in the context of different population sizes. The figure contains three graphs representing three different population sizes with the smallest on the left and the largest on the right. There are three percentages of veterans depicted on each graph denotated by red, blue, and green colored points and again the x-axis represents a ZIP code's increasing distance from a JROTC program. In the smallest population size, we can see a generally negative relationship between distance from a JROTC at all veteran population percentages. However, as ZIP code 18- to 24-Year-Old population sizes increases, as shown in the middle and rightmost graphs, higher veteran population sizes in conjunction with increasing distance from a JROTC, have a more dramatic association with the number of enlistments. Comparing the middle graph to the rightmost graph we can see the interactions become much larger with increasing population size.

Demonstrating the power of the interactions between distance from a JROTC and veteran population, Figure 7 shows changes in veteran population with a JROTC within the ZIP code (Factor 1). Looking at the rightmost graph with 18- to 24-Year-Old population size of 7500 and JROTC factor code 1, the veteran percentage size seems to have a large association with increases in number of enlistments. The model estimates (after transforming the response to counts) approximately 6 people enlisting at the red dot, 12 at the blue dot, and 20 at the green dot associating with progressively larger veteran population levels. Throughout the interaction plot, we see that the interactions seem to show a large association. It should be noted that the marginal difference along the veteran population is less significant when the JROTC location is within 5–10 miles from the ZIP code.

Figure 7. Interactions between JROTC Factor Code, Veteran population, and 18- to 24-Year-Old Population Sizes

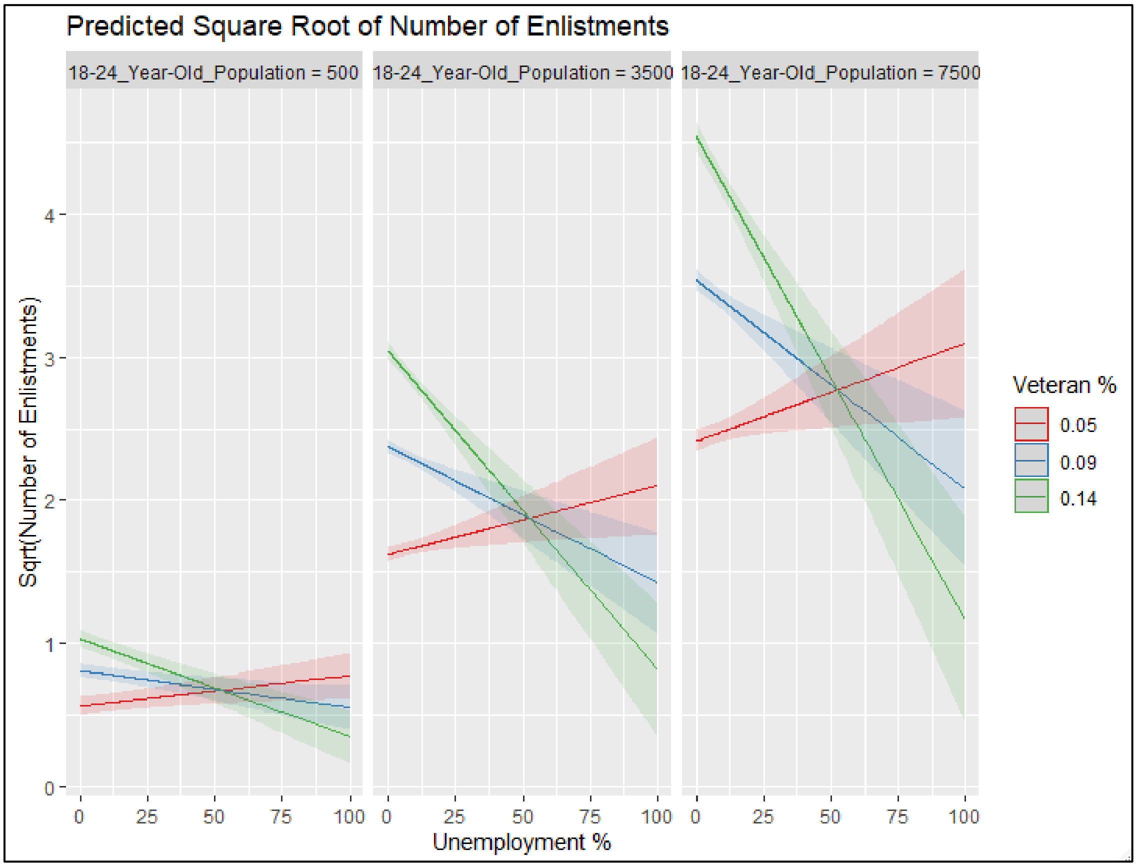


3. Interaction between Unemployment and Veteran Population

The interaction between Unemployment levels and Veteran populations provides another valuable insight. Figure 9 shows three different population sizes with the smallest on the left and largest on the right, with unemployment rates increasing from left to right on each graph. Veteran population percentages levels are shown as three levels; 5% as red dots, 9% as blue dots, and 14% as green dots. We can see at low levels of unemployment the size of the veteran population has a large association with the number of enlistments. However, as the unemployment rate increases and veteran populations increase the rate of enlistments drops considerably. This seems to indicate that in areas with smaller veteran populations, unemployment rates will generally correspond to an increasing number of enlistments. However, in areas with larger veteran populations, increasing levels of unemployment result in lowering the expected number of enlistments. We speculate that a possible reason for this is the perception that military service does not contribute to obtaining and maintaining employment after the service member leaves the service. In other words, people may see larger populations of veterans without a job and this may influence them to not join the Army. Additionally, Figure 8 demonstrates how larger populations dramatically shift the negative slope between veteran population and unemployment rates, and the number of enlistments.

An interesting question this raises is the importance of the military's efforts to assist veterans' transitions from the military. When developing the linear model, we determined that the economic status of the veteran population was not an important component on the number of enlistments. However, the factors we utilized did not separate veterans by age group. Examining unemployment rates or economic status of various age groups of veterans could provide additional insight.

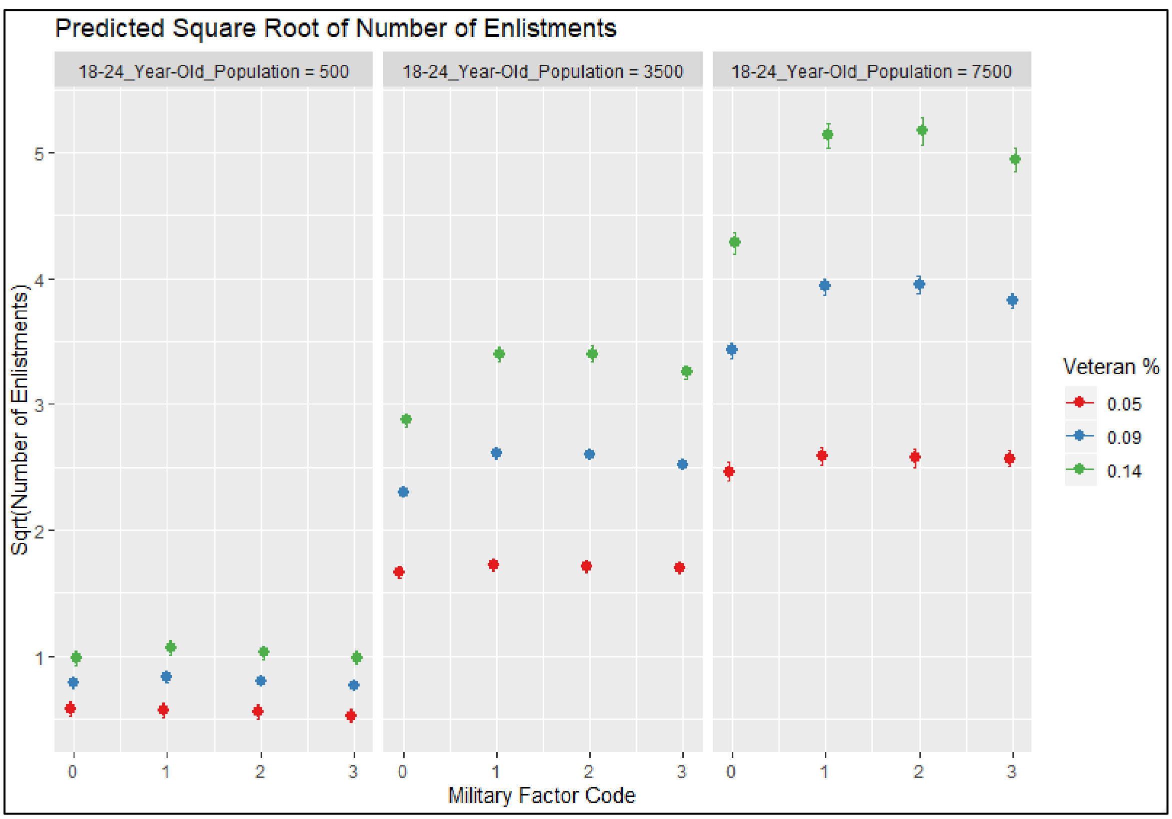
Figure 8. Interaction of Veteran Population and Unemployment Rates with Enlistments



4. Interaction between Veteran Population and distance from Military Installations

The idea that the size of a veteran population in proximity to military installations could play a role in the number of enlistments from an area is a fairly intuitive one. Indeed, Figure 9 shows how the two predictors are associated with enlistments. Figure 10 shows three different population sizes with the smallest on the left and largest on the right with veteran population percentages levels are illustrated as three levels with 5% as red dots, 9% as blue dots, and 14% as green dots. Recall from Table 4 that Military Factor Codes represent a ZIP code's distance from a military installation, with factor level 1 being within 10 miles from an installation and each increasing factor level being further away. A few insights we can gather from the interaction are that in areas with small populations surrounding a military installation, the effect of changes in veteran population is minimal. However, in larger population areas, large veteran populations have a much greater association. Looking at the rightmost graph with 18- to 24-Year-Old population size of 7500 and Military factor code 1, the veteran percentage size seems to have a large association with increases in number of enlistments. The model estimates (after transforming the response to counts) approximately 6 people enlisting at the red dot, 12 at the blue dot, and 20 at the green dot associating with progressively larger veteran population levels. As shown with larger population sizes, the effects of veteran populations and distance from military installations become even more pronounced.

Figure 9. Interactions between Veteran Population, 18- to 24-year-old Population, and DISTANCE from Military INSTALLATION



B. POISSON REGRESSION

While the Linear regression model provided interesting insights into the interactions between a number of the factors, the Poisson regression identified additional insights stemming from economic factors. A common belief is that individuals in lower income households will tend to opt for military service in an attempt to improve their station in life. Here we examine the association that the percentage of differing household income levels have on enlistments.

Figures 10–12 illustrate how greater percentages of household income levels influence the number of enlistments in a ZIP code at three different population sizes. In each figure, the red line depicts the smallest 18- to 24-Year-Old population size of 500 people, the blue represents a population size of 3500, and the green line represents 7500 people.

We can see in Figure 10 that there appears to be a negative association with an increasing percentage of households with an income of less than $25,000. We postulate that at the lowest income levels this trend exists because the financial resources may not exist in that area to support applicants from these areas to become “qualified” people to join the service. Common requirements for enlistment include no medical disqualifications, high school graduation (or GED), sufficiently high ASVAB score, and little to no history with law enforcement. Some of these requirements could be more difficult to meet in areas with significant poverty.

Figure 10. Effects of Percentage of Household Income Less Than $25,000 on Predicted Number of Enlistments

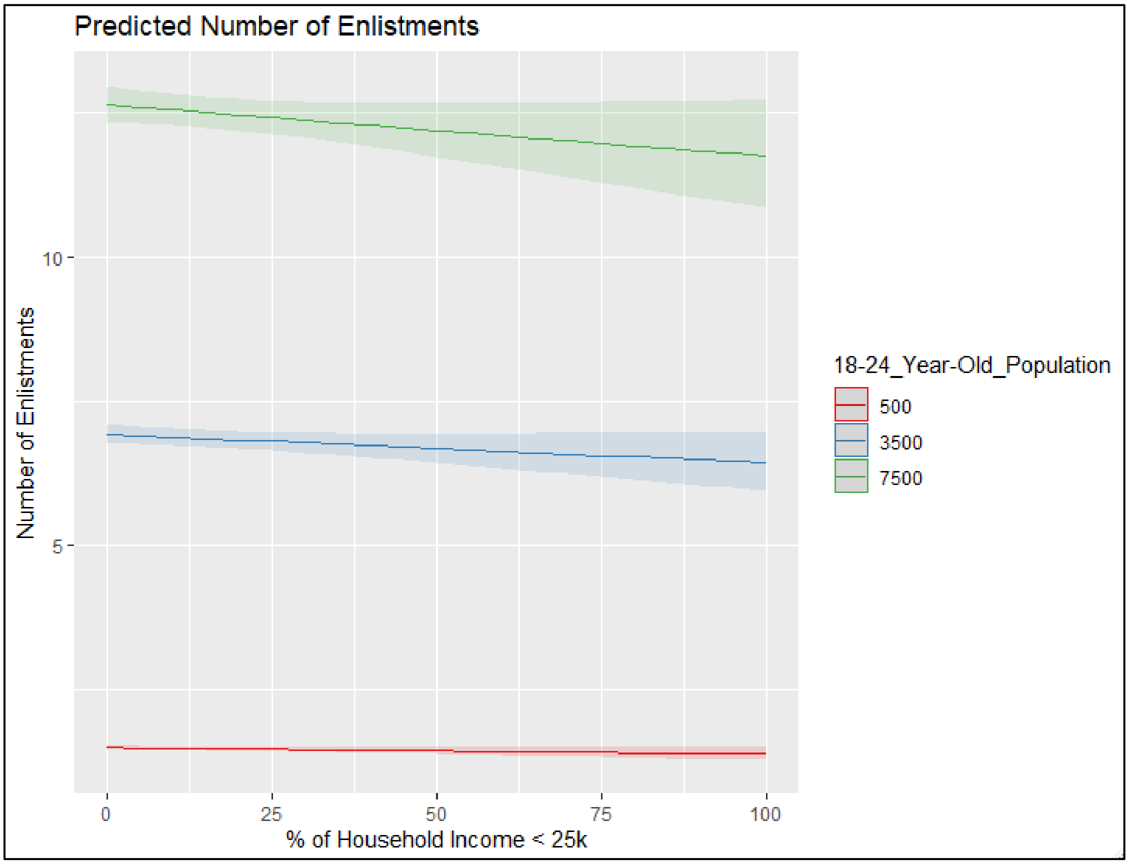


Figure 11 shows how increases in slightly higher income levels, household income levels between $25,000 and $50,000, change the association to be a positive one. We can also see that over the three population levels the slope of the positive association seems to be larger with larger population sizes. The trend seen in the graph would continue to support speculation that financial resources available in an area may influence the ability for that population to be qualified applicants and therefore lead to more enlistments.

Figure 11. Effects of Percentage of Household income Greater Than $25,000 but Less Than $50,000 on Predicted Number of Enlistments

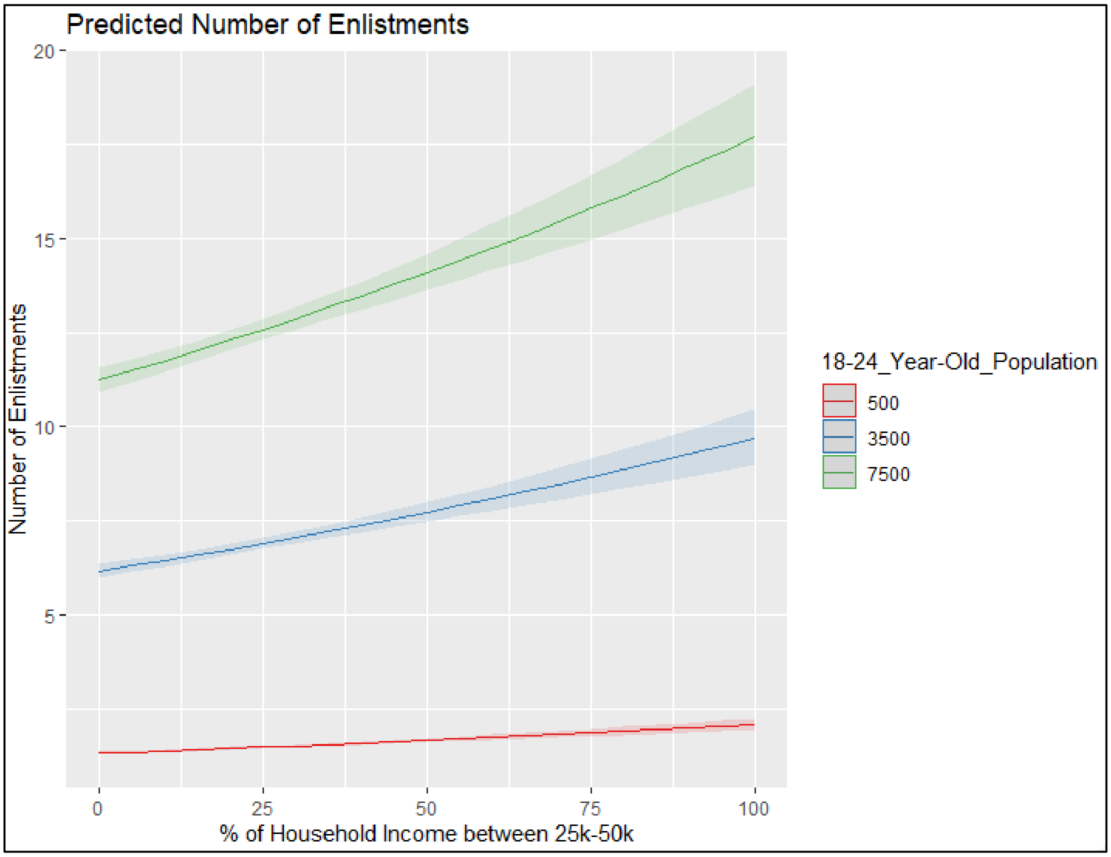
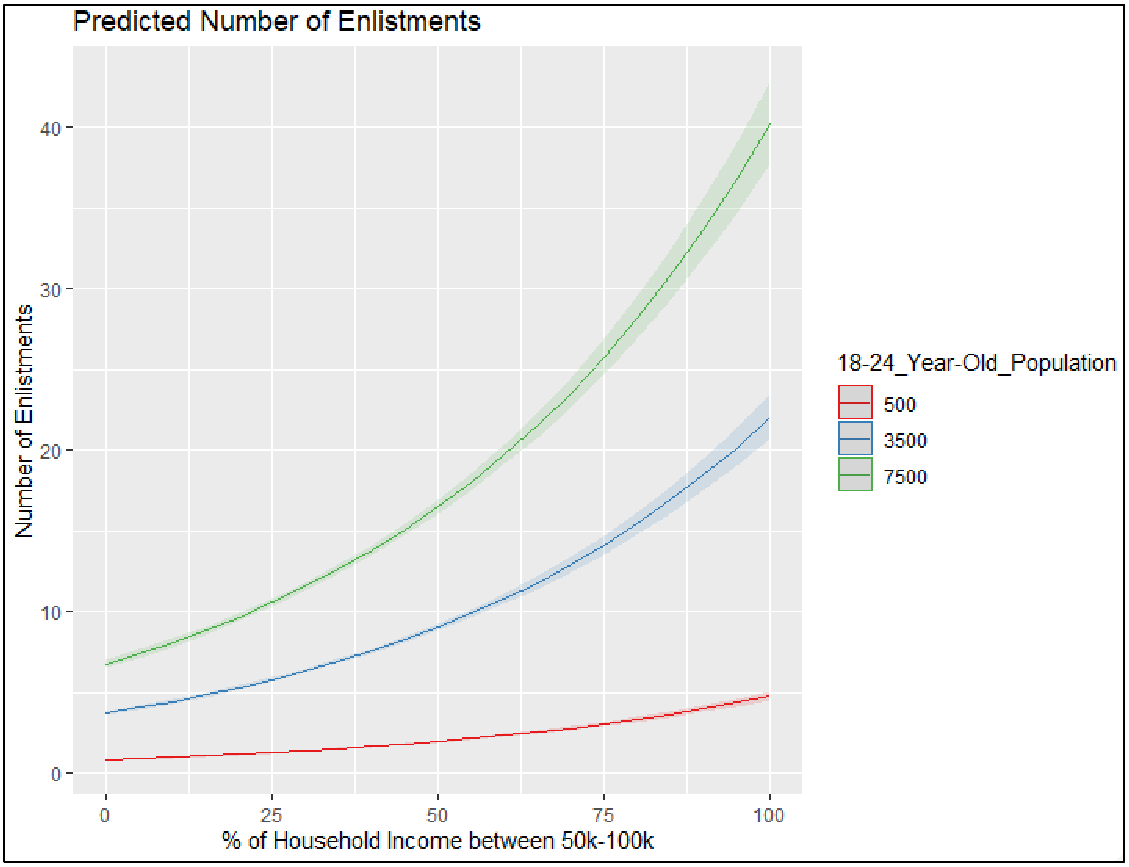


Figure 12 seems to further support our speculation that available financial resources in an area influences the number of enlistments. We can see that as the percentage of household incomes between $50,000 and $100,000, which we define as middle-class, increases the growth rate of enlistments become nearly quadratic with the upward bend of the curve increasing more dramatically with larger population sizes. Comparing across Figures 10 through 12, we can see that in the first two when the percentage of the households are low, they both seem to predict nearly the same number of enlistments. However, Figure 12 shows that when an area has very few middle-class households the predicted number of enlistments is much lower.

Figure 12. Effects of the Percentage of Household Income Greater Than $50,000 but Less Than $100,000 on Predicted Number of Enlistments



C. OPTIMAL ZIP CODE LOCATIONS

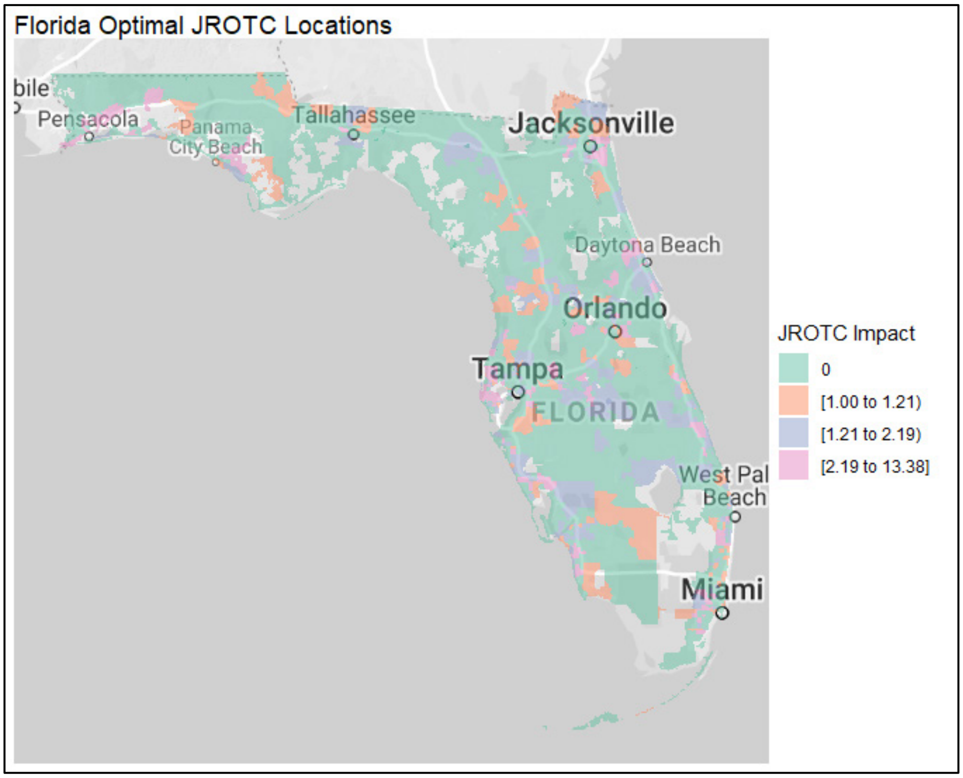
To determine the effects of establishing new JROTC programs in an area, we created an algorithm that determined the ZIP codes that are eligible for a program. A ZIP code was eligible if it did not already have an existing JROTC program. For each of these eligible ZIP codes, we identify all potentially impacted ZIP codes through the R package *ZipRadius* described in the methodology section (Ewing 2018). This gives us two groups of ZIP codes, one that establishes a JROTC within the ZIP code and the other group contains those ZIP codes that are within 15 miles of that ZIP code. In each iteration of the algorithm, we predicted the number of enlistments for both groups of ZIP codes using the original dataset. After calculating the predicted number of enlistments for each affected

ZIP code, we change the JROTC Factor Code value to 1 for the ZIP code that is establishing a JROTC program.

The other group of impacted ZIP codes identified would then have their JROTC Factor Codes adjusted if the Factor Code improved. For example, if a ZIP code originally had a Factor Code of 4, but due to the new JROTC location the affected ZIP would have a JROTC within 5 miles then the algorithm changed that Factor code to 2. However, if the JROTC factor code did not improve it was left unchanged. We then predicted the number of enlistments for each affected ZIP code using the same test data set with updated JROTC codes. We then ascertain a net improvement for each ZIP code by subtracting the original predicted number of enlistments from the new predictions.

This process created a total net improvement for each eligible ZIP code, which we then use to rank each ZIP code by its total overall effects on number of enlistments. Utilizing the R package *choroplethrZip*, we then overlay the various ZIP codes by state to graphically identify sections of the state that would show where establishing a JROTC would provide the most impact (Lamstein 2016). Figure 13 depicts a map of Florida with an overlay. Unsurprisingly, it highlights larger population areas like Tampa, Pensacola and other major cities as locations in which establishing a JROTC would have the most impact. Some of the areas are not colored as they represent areas that either have an established JROTC or areas where there are little to no population, e.g., the Everglades. Figures 13–16 all present the net improvement for each ZIP code for four different states, but the scale for each state is unique to that individual state.

Figure 13. Map of Florida Illustrating the Predicted Effect of Establishing a JROTC



Figures 13, 14, and 15 are further examples of establishing a JROTC within a state. Recall that one of the requirements for the JROTC program is to provide a more equitable distribution of the program around the country. Figure 13, which shows the state of map of Nebraska, provides an example for the Midwest states. Likely due to its smaller population, the map of Nebraska provides many fewer locations where adding a JROTC program would result in increased enlistments than the Florida one shows. New York, shown in Figure 14, shows a much larger range of impacts with the largest in New York City, Albany, and Syracuse. Washington State in Figure 15 demonstrates larger values around major cities but also extending north of Seattle towards Vancouver.

Figure 14. Map of Nebraska Illustrating the Predicted Effect of Establishing a JROTC

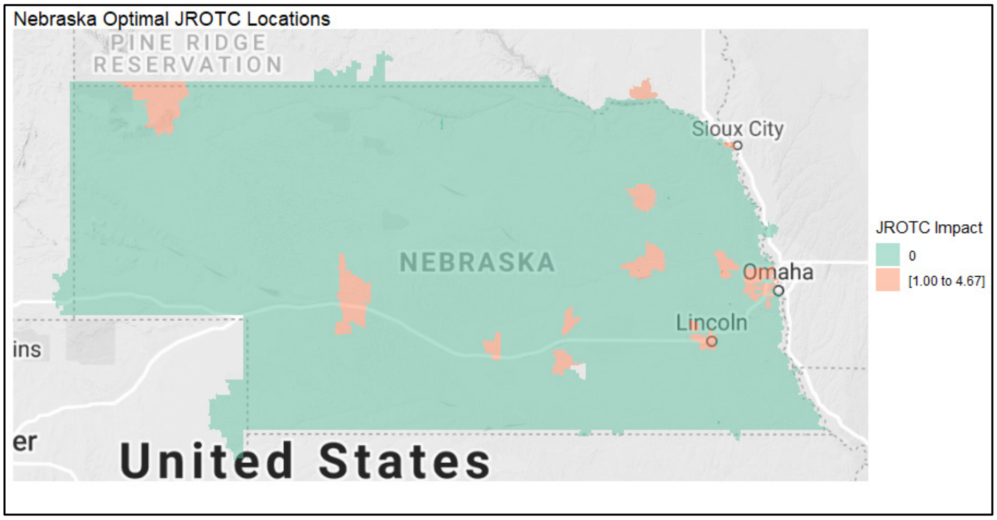


Figure 15. Map of New York Illustrating the Predicted Effect of Establishing a JROTC

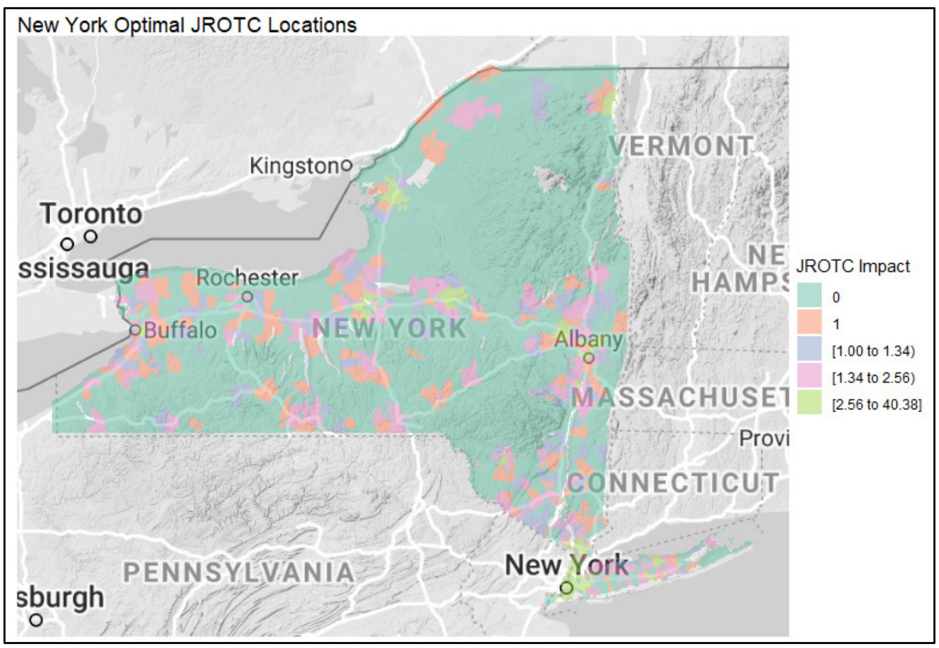
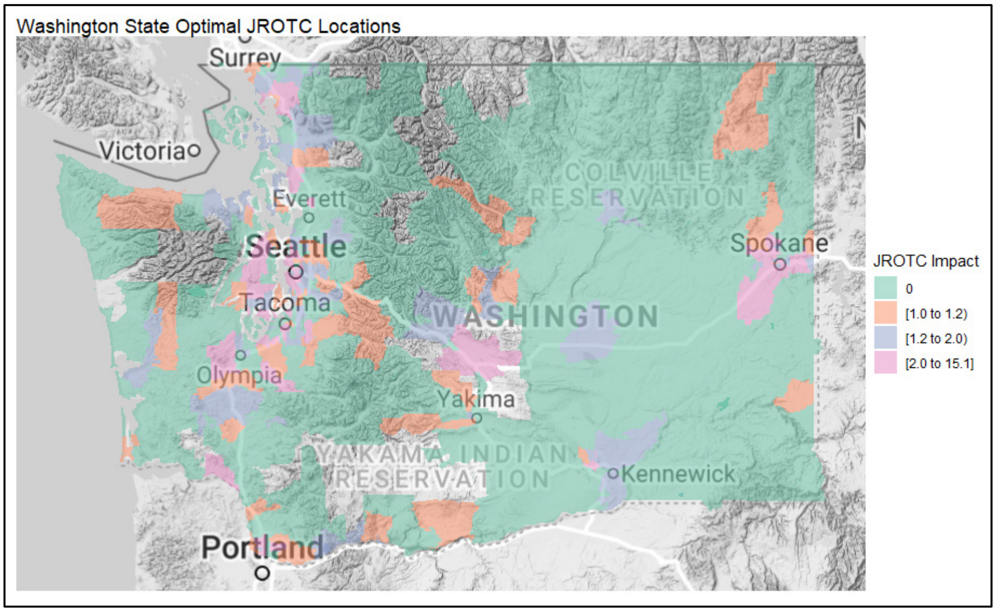


Figure 16. Map of Washington State Illustrating the Predicted Effect of Establishing a JROTC



VI. CONCLUSION AND FUTURE WORK

A. CONCLUSION

The primary focus of this thesis is to determine if there is an association between the presence of a JROTC program and the number of enlistments. Analysis conducted using a linear regression model from a dataset containing primarily economic and demographic factors indicates that the presence of a JROTC program is associated with a larger number of enlistments. However, the more significant finding indicates that while JROTC by itself does benefit the number of enlistments, the combination of a JROTC with larger veteran population provides an even more dramatic benefit. This agrees with conclusions from Goldberg (2018), which found that increased exposure to the military has positive impacts on enlistments. Two of the options he postulated that DOD could exercise to increase exposure to the military are JROTC programs and increase overall ASVAB testing in the general high school population (Goldberg 2018). Extending the use of the model developed in this analysis to identifying ZIP codes that provide the most benefit for recruiting efforts can be used to create an additional metric to measure the effectiveness of the JROTC programs. While recruiting is not the primary focus of the JROTC program, the program does seem to provide a positive association in generating enlistments both Cadet Command and senior DOD leadership can use to fully understand and develop a holistic approach to quantifying the overall Return on Investment of the program.

B. FUTURE RESEARCH

This study shows that JROTC does appear to have a positive association with recruiting. However, JROTC may provide other directly military benefits. It would be valuable to quantify the relationship between a recruit participating in JROTC and recruit "quality." Quality might include whether the service member completed his or her service commitment; promotion rates compared to non-JROTC service members, and the rate at which JROTC graduates re-enlist. It may be of value to include additional data on health and/or economic status of veterans, by age group, in this analysis. Other avenues of possible investigation include measuring the effectiveness of JROTC programs in providing higher quality leads for recruiting. Another avenue of research is studying the effects of the service specific bases, Active duty bases, Reserve bases, and National Guard Armories locations on enlistments. An important future research opportunity is to conduct this same analysis, instead looking at the change in enlistment rates per capita. Finally, the DOD could benefit from researching the effect on enlistments across all services, Reserves, and National Guard.

LIST OF REFERENCES

Army Junior ROTC (2020) Army Junior ROTC program overview. Accessed March 20, 2020. http://www.usarmyjrotc.com/general/program_overview.php

Buttrey S, Whitaker (2018) treeClust: An R package for tree-based clustering dissimilarities. Accessed February 13, 2020. R package version 1.1.7. https://CRAN.R-project.org/package=treeClust

Coumbe A, Harford L (1996) *U.S. Army Cadet Command The 10 Year History* U.S. Army Cadet Command. Fort Monroe, VA. Accessed December 20, 2019. https://apps.dtic.mil/dtic/tr/fulltext/u2/a317940.pdf

Dorminey D (2019a) Analysis of Junior Reserve Officer Training Corps (JROTC) U.S. Army Cadet Command.

Dorminey D (2019b) JROTC dataset provided to the author via email. September 25.

Department of Defense (2006) Junior Reserve Officers' Training Corps (JROTC) Program. DOD Instruction Number 1205.13. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/120513p.pdf

Glenn E (2019) acs: An R package to download, manipulate, and present American Community Survey and decennial data from the U.S. Census. Accessed November 7, 2019. R package version 2.1.4. https://CRAN.R-project.org/package=acs

Goldberg M, Cheng K, Huff N, Kimko D, Saizan A (2018) Geographic diversity in military recruiting. Report IDA Paper P-9079, Institute for Defense Analyses, Alexandria, VA. https://www.ida.org/-/media/feature/publications/g/ge/geographic-diversity-in-military-recruiting/d-9079.ashx

Goldman C, Schweig J, Buenaventura M, Wright C (2017) Geographic and demographic representativeness of Junior Reserve Officer Training Corps. Report RR-1712, RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR1712.html

Ewing S (2020) ZipRadius: An R package that creates a data frame of U.S. Zip Codes in given radius from a given U.S. Zip Code. Accessed January 7, 2020. R package version 1.0.1. https://CRAN.R-project.org/package=ZipRadius

Fulton B (2016) Determining market categorization of United States Zip Codes for purposes of Army recruiting. Master's thesis, Naval Postgraduate School. Accessed December 10, 2019. https://calhoun.nps.edu/bitstream/handle/10945/49463/16Jun_Fulton_Brandon.pdf?sequence=1&isAllowed=y

Kahle D, Wickham H (2020) ggmap: An R package that creates spatial visualization with ggplot2. R package version 3.0. Accessed January 10, 2020, https://CRAN.R-project.org/package=ggmap

Jackman S (2020) pscl: Political Science Computational Laboratory. R package version 1.5.5. Accessed February 13, 2020. https://CRAN.R-project.org/package=pscl

Kamarack K, Arriagea X, (2020) Defense Primer: Junior Reserve Officer Training Corps (CRS Report No. IF11313) Accessed from Congressional Research Service website. https://fas.org/sgp/crs/natsec/IF11313.pdf

Lamstein A (2016) choroplethrZip: An R package with shapefile, metadata, and visualization functions for U.S. Zip Code Tabulated Areas (ZCTAs). Accessed March 4, 2020. R package version 1.5. https://github.com/arilamstein/choroplethrZip

Myers M (2018) Army Times: To draw more soldiers, the Army wants more recruiters, bigger budgets and a better slogan (November 7). https://www.armytimes.com/news/your-army/2018/11/07/to-draw-more-soldiers-the-army-wants-more-recruiters-bigger-budgets-and-a-better-slogan/

National Center for Education Statistics (2020) Fast Facts: Title I. Accessed April 29, 2020, https://nces.ed.gov/fastfacts/display.asp?id=158.

National Center for Education Statistics (2019) Postsecondary data tables. Accessed December 5. https://nces.ed.gov/datalab/TablesLibrary

Opendatasoft (2020) Military bases, published as part of U.S. Department of Transportation/Bureau of Transportation Statistics National Transportation Atlas database. Accessed March 20, 2020, https://public.opendatasoft.com/explore/dataset/military-bases/information/?refine.country=United+States&refine.oper_stat=Active

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Rodriguez, German (2016) *Poisson Models for Count Data. Lecture Notes, Generalized Linear Statistical Models*, Princeton University, Princeton, NJ. https://data.princeton.edu/wws509/notes/c4.pdf

U.S. Army Recruiting Command (2020) Aggregate number of enlistments by age and ZIP code. November 15.

United States Census Bureau (2020) Demographic datasets 2015–2017 American Community Survey 5-year estimates. Accessed December 10, 2019.

United States Census Bureau (2015) ZIP Code Tabulation Areas (ZCTAS) Brochure February 2015. Accessed April 3, 2020. https://www2.census.gov/geo/pdfs/education/brochures/ZCTAs.pdf

United States Postal Service Office of Inspector General (2013) The untold story of the ZIP code. Accessed March 13, 2020, https://www.uspsoig.gov/sites/default/files/document-library-files/2015/rarc-wp-13-006_0.pdf

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