Best Available Copy
THESIS

AN ANALYSIS OF THE COAST GUARD ENLISTED ATTRITION

by

Laureano Enrique Oñate Rubiano

September 1993

Thesis Advisor: So Young Sohn

Approved for public release; distribution is unlimited.
In this thesis, survival analysis is used to study US Coast Guard enlisted attrition behavior in terms of individual personnel characteristics such as sex, marital status, race, paygrade and rating. Results obtained based on 8 years of historical data from FY83 to FY90 are as follows: males and married individuals have higher survival probabilities than their counterparts, respectively; paygrades E-1 to E-5 have higher attrition than paygrades E-6 to E-9; American Indians have the highest attrition and Asian members have the highest survival probabilities; rating 170 (Gunner's Mate) has the highest attrition over all ratings followed by rating 180 (Fire Control Technician); the rating with the highest survival probability is 570 (Aviation Machinist's Mate); a decreasing trend in attrition was found during the last 4 years of the observation period; it was also observed that there was significantly high attrition at the end of the four years service contract and when the enlisted member reached twenty years of service. Additionally, this thesis provides the a regression model in order to predict monthly enlisted attrition figures. Significant predictors selected are the prior month's attrition, the number of enlistments four years prior and the current unemployment rate. The selected regression model explains almost 97% of the total variation of monthly attrition. It turns out to perform better than the current method used by the CG.
AN ANALYSIS OF THE COAST GUARD ENLISTED ATTRITION

by

Laureano Enrique Oñate Rubiano
Lieutenant Commander, Colombian Navy
B.S., Escuela Naval de Cadetes Cartagena Colombia, 1987

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN OPERATION ANALYSIS

from the

NAVAL POSTGRADUATE SCHOOL
September 1993

Author:
Laureano Enrique Oñate Rubiano

Approved by:
So Young Sohn, Thesis Advisor

Robert R. Read, Second Reader
Peter Purdue, Chairman
Department of Operation Research
ABSTRACT

In this thesis, survival analysis is used to study US Coast Guard enlisted attrition behavior in terms of individual personnel characteristics such as sex, marital status, race, paygrade and rating. Results obtained based on 8 years of historical data from FY83 to FY90 are as follows: males and married individuals have higher survival probabilities than their counterparts, respectively; paygrades E-1 to E-5 have higher attrition than paygrades E-6 to E-9; American Indians have the highest attrition and Asian members have the highest survival probabilities; rating 170 (Gunner’s Mate) has the highest attrition over all ratings followed by rating 180 (Fire Control Technician); the rating with the highest survival probability is 570 (Aviation Machinist’s Mate); a decreasing trend in attrition was found during the last 4 years of the observation period; it was also observed that there was significantly high attrition at the end of the four years service contract and when the enlisted member reaches twenty years of service.

Additionally, this thesis provides a regression model in order to predict monthly enlisted attrition figures. Significant predictors selected are the prior month’s attrition, the number of enlistments four years prior and the current unemployment rate. The selected regression model explains almost 97% of the total variation of monthly attrition. It turns out to perform better than the current method used by the CG.
# TABLE OF CONTENTS

I. INTRODUCTION .............................................. 1
   A. PROBLEM DESCRIPTION ................................... 1
   B. OBJECTIVES ............................................. 2
   C. PERSONNEL BACKGROUND ................................. 2
   D. SCOPE OF THE THESIS ................................... 5

II. DATA OVERVIEW .............................................. 7
   A. POPULATION .............................................. 7
   B. SUMMARY OF PERSONNEL ................................. 10

III. SURVIVAL ANALYSIS ......................................... 14
   A. SURVIVAL FUNCTIONS ...................................... 14
      1. Enlisted attrition by sex. ........................... 16
      2. Enlisted attrition by marital status. ............. 17
      3. Enlisted attrition by race .......................... 18
      4. Enlisted attrition by paygrade ...................... 19
      5. Enlisted attrition by rating ......................... 20
   B. SUMMARY ................................................ 20

IV PREDICTION MODEL FOR MONTHLY ATTRITION ................. 23
   A. MULTIPLE REGRESSION MODEL ............................ 23
   B. RESULTS ................................................ 27
LIST OF FIGURES

Figure 1. Enlisted attrition by month each year. . . . 8
Figure 2. First quarter FY84 fixed. . . . . . . . . . . 9
Figure 3. Enlisted attrition by sex. . . . . . . . . . . 16
Figure 4. Enlisted attrition by marital status. . . . . . 17
Figure 5. Enlisted attrition by race. . . . . . . . . . . 18
Figure 6. Enlisted attrition by paygrade. . . . . . . . 19
Figure 7. Enlisted attrition by rating. . . . . . . . . . 21
Figure 8. Validation of model 2 and current CG method. 25
Executive Summary

One of the responsibilities of the US Coast Guard Personnel Workforce Planning Office is to forecast enlisted monthly attrition. In order to forecast enlisted monthly attrition, it is necessary to analyze attrition behavior. The current method to forecast the number of enlisted attrition is based on the mean of the past eight years attrition figures of the corresponding month. This method facilitates simple implementation but it does not utilize other useful information such as economy, attrition behavior and individual characteristics associated with a specific kind of enlisted member.

In this thesis, first, survival analysis was applied to investigate the attrition behavior of the US Coast Guard enlisted in terms of their individual characteristics such as sex, marital status, race, paygrade and rating. Results obtained based on the past eight years data from October of 1982 to September of 1990 are as follows: males and married individuals have higher survival probabilities than their counterparts, respectively; paygrades E-1 to E-5 have higher attrition than paygrades E-6 to E-9; American Indians have the highest attrition and the Asian members the highest survival functions; rating 170 (Gunner’s Mate) has the highest attrition over all ratings followed by rating 180.
(Fire Control Technician); the rating with the highest survival probability is 570 (Aviation Machinist’s Mate); there was a decreasing trend in attrition in the last 4 years of the observation period. The common feature of the survival functions was the significant fall at the end of the four years service contract and when the enlisted member reach twenty years of service.

Secondly, based on the observations made in the survival analysis a candidate set of predictors was selected to fit a regression model for enlisted monthly attrition. The model finally selected contains the prior month’s attrition, the number of enlistments four years prior and the current unemployment rate as significant predictors. The selected regression model explains almost 97% of the total variation of monthly attrition. Performance of the selected model is better than the current method used by US Coast Guard in terms of the mean squared error and the mean relative error.
I. INTRODUCTION

A. PROBLEM DESCRIPTION

One of the responsibilities of the United States Coast Guard (CG) Personnel Workforce Planning (PWP) Office, located in Washington D.C., is to forecast personnel stocks, promotion requirements, monthly attrition and recruitment needs. The current method to forecast the number of enlisted attrition by paygrade per month is based on the mean of the past previous eight years attrition figures. For the corresponding month this method facilitates simple implementation. However, it does not utilize other useful information such as attrition behavior and individual characteristics associated with a specific kind of enlisted member. The CG Office of Personnel and Training maintains enlisted personnel records in addition to information concerning the main reasons for enlisted attrition. Recently a study was written [Ref. 1], examining enlisted attrition behavior and developing a model that projects these attrition figures. The study lacked of temporal stability, as it used personnel data for only one year (FY91). The conclusions and modelling results were necessarily not very reliable.
B. OBJECTIVES

This thesis uses eight years of historical data from FY83 through FY90 to obtain the survival functions of the individual personnel characteristics for survival analysis. Additionally, it applies regression models to predict monthly attrition figures and compares the performance of the resulting prediction model to those previously in use.

C. PERSONNEL BACKGROUND

A summary from USGC LT Douglas Allen Blakemore’s thesis is used to describe the personnel background of the Coast Guard enlisted personnel structure.

The Coast Guard enlisted personnel structure is composed of nine paygrades from E-1 to E-9. The paygrade E-1 contains recruits attending Coast Guard basic training school; paygrade E-2 is composed of enlisted personnel who have completed basic training and have been assigned to active duty commands in preparation for attending a CG specialty school. Paygrades E-3 through E-9 are divided into 24 active duty military occupational skills (MOS) or subspecialities. Of these, only 22 are of concern since one (Musician), has an extremely small attrition rate and the other (Sonar Technician), no longer exists. The following list contains MOS; their respective CG abbreviation for subspecialty name and the rating code. It will be used throughout this thesis to calculate survival probabilities.
<table>
<thead>
<tr>
<th>MOS Subspecialty</th>
<th>Rating</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Aviation Machinist’s Mate</td>
<td>AM</td>
<td>570</td>
</tr>
<tr>
<td>- Aviation Electrician’s Mate</td>
<td>AE</td>
<td>560</td>
</tr>
<tr>
<td>- Aviation Damage Controlman</td>
<td>AD</td>
<td>520</td>
</tr>
<tr>
<td>- Aviation Survivalman</td>
<td>ASM</td>
<td>530</td>
</tr>
<tr>
<td>- Aviation Electronic Technician</td>
<td>AT</td>
<td>550</td>
</tr>
<tr>
<td>- Boatwain’s Mate</td>
<td>BM</td>
<td>100</td>
</tr>
<tr>
<td>- Damage Controlman</td>
<td>DC</td>
<td>210</td>
</tr>
<tr>
<td>- Electrician’s Mate</td>
<td>EM</td>
<td>270</td>
</tr>
<tr>
<td>- Electronics Technician</td>
<td>ET</td>
<td>240</td>
</tr>
<tr>
<td>- Fire Control Technician</td>
<td>FT</td>
<td>180</td>
</tr>
<tr>
<td>- Gunner’s Mate</td>
<td>GM</td>
<td>170</td>
</tr>
<tr>
<td>- Health Service Technician</td>
<td>HS</td>
<td>870</td>
</tr>
<tr>
<td>- Machinery Technician</td>
<td>MK</td>
<td>200</td>
</tr>
<tr>
<td>- Marine Science Technician</td>
<td>MST</td>
<td>790</td>
</tr>
<tr>
<td>- Public Affairs Specialist</td>
<td>PA</td>
<td>340</td>
</tr>
<tr>
<td>- Quartermaster</td>
<td>QM</td>
<td>110</td>
</tr>
<tr>
<td>- Radarman</td>
<td>RD</td>
<td>130</td>
</tr>
<tr>
<td>- Radioman</td>
<td>RM</td>
<td>350</td>
</tr>
<tr>
<td>- Storekeeper</td>
<td>SK</td>
<td>420</td>
</tr>
<tr>
<td>- Subsistence Specialist</td>
<td>SS</td>
<td>500</td>
</tr>
<tr>
<td>- Telephone Technician</td>
<td>TT</td>
<td>280</td>
</tr>
<tr>
<td>- Yeoman</td>
<td>YN</td>
<td>360</td>
</tr>
</tbody>
</table>

An enlisted man (EM) receives an MOS after completing a specialty school or completing an intensive on-the-job
training program. Both methods are usually administered while the EM is in the E-3 paygrade. An E-3 who has not obtained an MOS is called non-rated; E-3’s through E-9’s hold only one MOS at a time and usually maintain that MOS throughout their careers.

Promotion to the next highest paygrade is determined by MOS and dictated by the needs of the CG. Promotions to paygrade E-5 through E-9 are vacancy driven while those to paygrades E-3 and E-4 occur on qualification for advancement. Promotion to paygrade E-2 occurs upon completion of basic training.

Enlisted personnel sign service obligation contracts (enlistments) that require the individual to serve in the CG for a pre-determined number of years, usually four years. (In the past, there have been two year enlistments but these have been terminated.) Upon completion of an enlistment and upon approval of the CG, a person may sign a new contract (re-enlist), or separate from the CG. Enlisted attrition generally occurs due to:

- Retirement - After 20 years of active duty service.
- Non-re-enlistments - An EM chooses not to reenlist or the CG chooses not to reenlist the individual.
- Administrative reasons - An EM may depart the CG prior to the end of his/her contract enlistment -" for the convenience of the government ".
- Selection to an Officer program.
- Death or disability.
This study will concentrate on the first four reasons listed above and will use the word "separation" referring to the EM leaving active duty service.

D. SCOPE OF THE THESIS

The first goal of this study is to develop survival functions for USCG EM personnel, in which survival analysis techniques can be used to analyze data on the length of time an EM remained in CG. This technique takes on different names, depending on the particular application at hand.

Recently the term, "event history analysis", has been used by social scientists to describe applications for the analysis of the length of time it takes an employee to retire or resign from a given job [Ref. 2].

Survival analysis is a method for describing the distribution of the length of time for a given event, such as the termination of the service in the CG. One way to perform survival analysis is to construct a histogram for the length of time that individuals spend in service. Alternatively, one can use the length of service time as a dependent variable and determine if it can be predicted by variables such as marital status, race, gender or military occupational skill.

The main tool used in the survival analysis is the survival function calculated from the data using the frequency histogram in which for any given time t, the area under the curve to the left of t is the proportion of individuals in the
population who separate from the CG to time t. A common feature of survival data is the presence of right censored observations due either to withdrawal of experimental units or termination of the experiment. For such observations it is only known that the lifetime exceeded the given value. The exact lifetime remains unknown. Such data cannot be analyzed by ignoring the censored observations because, among other considerations, the longer-lived observations are generally more likely to be censored. The survival distribution function evaluated at t is the probability that an observation from the population will have a lifetime exceeding t, that is $S(t) = \text{Prob}(T > t)$. Their estimators are called product limit estimators. [Ref 2].

An analysis with the entire data set will be elaborated, in order to know some specific trends, behaviors and predominance of the individual characteristics of the data under study.

A regression model will be developed to forecast monthly attrition and to establish relationships between explanatory variables and the attrition.
II. DATA OVERVIEW

A. POPULATION

The USCG Office of Personnel and Training located in Washington D.C., provided the data for this study. The data set contained 27,2160 individual personal records for all CG enlisted personnel for the observation period from FY83 to FY90. The fields contained in each record are: Personnel Identification Number (PID), a generic number which identifies one individual in the CG; rating, according to the military occupational skill; paygrade, from E-1 to E-9; sex, males and females; minority designator, race of the individual; marital status, married and single; date of entrance into the CG; date of separation from the CG; GAOCD, is a separation designator code for CG enlisted personnel; censor, which indicates whether the person remained in the CG or not. The data contains a record for each enlisted member per year of service in the CG, in other words, if a person remains in service for 8 years there exists 8 records for that individual. These multiple records were collapsed as one record per person, without loss of information. After this process, there were 50,036 records from which 29,405 belonged to people who left active duty service during the observation period and 20631 which belonged to censored active duty CG
members, at the end of FY90 (30 Sep/90). This data includes cases where the CG allowed a member to separate from the CG and then rejoin in the future at the member former paygrade.

![Enlisted Attrition by Month Each Year](image)

**Figure 1.** Enlisted attrition by month each year.

Based on this data set the number of months a person spent in the CG can be calculated, to analyze total time in service. In addition, censored information can be obtained. Figure 1 shows the monthly attrition for each year from FY83 to FY90. The attrition for Dec/83 appears to be an outlier since the attrition values for the other months that remain are no
are no greater than 650. However, this value is preceded by the two smallest attrition values in the sample (25 and 30) and they both are under the next minimum value of attrition (182).

![Enlisted Attrition by Month Each Year](image)

**Figure 2.** First quarter FY84 fixed.

In the future, analysis will conducted on the first quarter of FY84 as an average of the three attrition values corresponding to Oct/83 (25), Nov/83 (30) and Dec/83 (1,339). As can be seen in Figure 2, the new values for FY84 are in the ranges of all other fiscal years values.
B. SUMMARY OF PERSONNEL

Table 2 provides a comprehensive view of the stocks of personnel by paygrade of enlisted active duty (AC) at 30 Sep/90, and the total of enlisted members who separated (RE) during the observation period from starting FY83 to the end of FY90. Of note is the fact that 22% of the total data is comprised of those members who are in the E-2 paygrade and the highest amount of attrition in paygrade E-4.

Table 2. Active Duty Enlisted at 30 Sep/90 and Separated Enlisted Personnel from FY83 to FY90

<table>
<thead>
<tr>
<th>PAYGRADE</th>
<th>ACTIVE DUTY</th>
<th>SEPARATED</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-1</td>
<td>1781</td>
<td>4778</td>
<td>6559</td>
</tr>
<tr>
<td>E-2</td>
<td>6894</td>
<td>4024</td>
<td>10918</td>
</tr>
<tr>
<td>E-3</td>
<td>4116</td>
<td>4918</td>
<td>9034</td>
</tr>
<tr>
<td>E-4</td>
<td>1893</td>
<td>7619</td>
<td>9512</td>
</tr>
<tr>
<td>E-5</td>
<td>2003</td>
<td>4252</td>
<td>6255</td>
</tr>
<tr>
<td>E-6</td>
<td>2548</td>
<td>1880</td>
<td>4428</td>
</tr>
<tr>
<td>E-7</td>
<td>1199</td>
<td>1084</td>
<td>2283</td>
</tr>
<tr>
<td>E-8</td>
<td>137</td>
<td>441</td>
<td>578</td>
</tr>
<tr>
<td>E-9</td>
<td>60</td>
<td>409</td>
<td>469</td>
</tr>
<tr>
<td>TOTAL</td>
<td>20631</td>
<td>29405</td>
<td>50036</td>
</tr>
</tbody>
</table>
Table 3. Personnel by sex and Marital Status

<table>
<thead>
<tr>
<th>PAYGRADE</th>
<th>SEX</th>
<th>MARITAL STATUS</th>
<th>TOTALS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MALE</td>
<td>FEMALE</td>
<td>SINGLE</td>
</tr>
<tr>
<td>E-1</td>
<td>AC</td>
<td>1592</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>4009</td>
<td>769</td>
</tr>
<tr>
<td>E-2</td>
<td>AC</td>
<td>6200</td>
<td>694</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>3359</td>
<td>665</td>
</tr>
<tr>
<td>E-3</td>
<td>AC</td>
<td>3808</td>
<td>308</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>4328</td>
<td>590</td>
</tr>
<tr>
<td>E-4</td>
<td>AC</td>
<td>1765</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>6947</td>
<td>672</td>
</tr>
<tr>
<td>E-5</td>
<td>AC</td>
<td>1895</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>3888</td>
<td>364</td>
</tr>
<tr>
<td>E-6</td>
<td>AC</td>
<td>2497</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>1802</td>
<td>78</td>
</tr>
<tr>
<td>E-7</td>
<td>AC</td>
<td>1198</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>1083</td>
<td>1</td>
</tr>
<tr>
<td>E-8</td>
<td>AC</td>
<td>136</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>441</td>
<td>0</td>
</tr>
<tr>
<td>E-9</td>
<td>AC</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>409</td>
<td>0</td>
</tr>
<tr>
<td>TOT</td>
<td>AC</td>
<td>19151</td>
<td>1480</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>26266</td>
<td>3139</td>
</tr>
</tbody>
</table>

Table 3, present the stocks of personnel by sex and marital status. As can be observed, males, and single members are the dominant groups in the CG; females have reached paygrade E-8; married members (4,813) compared to single (16,526), left from the CG in low Paygrades (E-1 to E-4), in lower proportion compared with paygrades E-5 to E-9 (married (4,387), single (3,679)); there are small amount of single members in paygrades E-8 and E-9 in active duty.
Table 4 contains the enlisted stocks classified by race. The findings observed here are: Caucasian has the largest stocks followed by Blacks, Hispanic, American Indians and Asian.

Table 4. Active Duty and Retired personnel by race

<table>
<thead>
<tr>
<th>PAYGRADE</th>
<th>RACE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BLACK</td>
<td>HISPA</td>
</tr>
<tr>
<td>E-1</td>
<td>AC</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>563</td>
</tr>
<tr>
<td>E-2</td>
<td>AC</td>
<td>412</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>534</td>
</tr>
<tr>
<td>E-3</td>
<td>AC</td>
<td>306</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>656</td>
</tr>
<tr>
<td>E-4</td>
<td>AC</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>691</td>
</tr>
<tr>
<td>E-5</td>
<td>AC</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>309</td>
</tr>
<tr>
<td>E-6</td>
<td>AC</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>107</td>
</tr>
<tr>
<td>E-7</td>
<td>AC</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>36</td>
</tr>
<tr>
<td>E-8</td>
<td>AC</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>10</td>
</tr>
<tr>
<td>E-9</td>
<td>AC</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>8</td>
</tr>
<tr>
<td>TOT</td>
<td>AC</td>
<td>1442</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>2914</td>
</tr>
</tbody>
</table>

During the eight years of observation American Indians have reached only paygrade E-7. Blacks, Hispanics, American Indians and Asian together are only the 13% of the population.
Table 5, classifies each paygrade by rating. The non-rated enlisted (MOS 150 and 320) comprise 49% of the total population, the rating with the greatest amount of enlisted is 200 (Machinery technician) with 9%, followed by rating 100 (Boatwain's Mate) with 6.6%. The rating least populated is 180 (Fire Control Technician) with 114 enlisted (0.2%).

Table 5. Enlisted personnel by MOS

<table>
<thead>
<tr>
<th>RATE</th>
<th>E-1</th>
<th>E-2</th>
<th>E-3</th>
<th>E-4</th>
<th>E-5</th>
<th>E-6</th>
<th>E-7</th>
<th>E-8</th>
<th>E-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0</td>
<td>25</td>
<td>131</td>
<td>1122</td>
<td>681</td>
<td>696</td>
<td>450</td>
<td>80</td>
<td>119</td>
</tr>
<tr>
<td>110</td>
<td>0</td>
<td>13</td>
<td>62</td>
<td>407</td>
<td>266</td>
<td>233</td>
<td>134</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>130</td>
<td>0</td>
<td>9</td>
<td>29</td>
<td>187</td>
<td>86</td>
<td>84</td>
<td>41</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>150</td>
<td>655</td>
<td>8083</td>
<td>6002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>170</td>
<td>0</td>
<td>8</td>
<td>24</td>
<td>240</td>
<td>85</td>
<td>83</td>
<td>38</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>180</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>43</td>
<td>33</td>
<td>12</td>
<td>12</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>72</td>
<td>300</td>
<td>1647</td>
<td>1137</td>
<td>771</td>
<td>426</td>
<td>83</td>
<td>59</td>
</tr>
<tr>
<td>210</td>
<td>0</td>
<td>9</td>
<td>66</td>
<td>418</td>
<td>287</td>
<td>188</td>
<td>84</td>
<td>19</td>
<td>11</td>
</tr>
<tr>
<td>240</td>
<td>0</td>
<td>16</td>
<td>65</td>
<td>790</td>
<td>517</td>
<td>327</td>
<td>158</td>
<td>46</td>
<td>25</td>
</tr>
<tr>
<td>270</td>
<td>0</td>
<td>23</td>
<td>86</td>
<td>633</td>
<td>193</td>
<td>144</td>
<td>88</td>
<td>37</td>
<td>31</td>
</tr>
<tr>
<td>280</td>
<td>0</td>
<td>2</td>
<td>10</td>
<td>115</td>
<td>85</td>
<td>46</td>
<td>29</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>320</td>
<td>0</td>
<td>2472</td>
<td>1650</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>340</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>43</td>
<td>31</td>
<td>23</td>
<td>13</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>350</td>
<td>0</td>
<td>28</td>
<td>112</td>
<td>590</td>
<td>419</td>
<td>284</td>
<td>118</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>360</td>
<td>0</td>
<td>6</td>
<td>65</td>
<td>698</td>
<td>545</td>
<td>358</td>
<td>174</td>
<td>44</td>
<td>42</td>
</tr>
<tr>
<td>420</td>
<td>0</td>
<td>11</td>
<td>75</td>
<td>510</td>
<td>460</td>
<td>255</td>
<td>87</td>
<td>32</td>
<td>21</td>
</tr>
<tr>
<td>500</td>
<td>0</td>
<td>121</td>
<td>230</td>
<td>822</td>
<td>361</td>
<td>262</td>
<td>423</td>
<td>35</td>
<td>26</td>
</tr>
<tr>
<td>520</td>
<td>0</td>
<td>4</td>
<td>24</td>
<td>303</td>
<td>201</td>
<td>131</td>
<td>59</td>
<td>39</td>
<td>22</td>
</tr>
<tr>
<td>530</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>63</td>
<td>44</td>
<td>29</td>
<td>12</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>550</td>
<td>0</td>
<td>3</td>
<td>24</td>
<td>184</td>
<td>223</td>
<td>136</td>
<td>56</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>560</td>
<td>0</td>
<td>2</td>
<td>17</td>
<td>164</td>
<td>117</td>
<td>78</td>
<td>39</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>570</td>
<td>0</td>
<td>1</td>
<td>18</td>
<td>198</td>
<td>124</td>
<td>88</td>
<td>40</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>790</td>
<td>0</td>
<td>1</td>
<td>14</td>
<td>65</td>
<td>59</td>
<td>35</td>
<td>20</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>870</td>
<td>0</td>
<td>5</td>
<td>18</td>
<td>270</td>
<td>301</td>
<td>165</td>
<td>82</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>TOT</td>
<td>6559</td>
<td>10918</td>
<td>9034</td>
<td>9512</td>
<td>6255</td>
<td>4428</td>
<td>2283</td>
<td>578</td>
<td>469</td>
</tr>
</tbody>
</table>
III. SURVIVAL ANALYSIS

This chapter presents methods for further quantifying the probability that an enlisted member will separate from the CG. The time reference is established as the number of months an enlisted member served in the CG. This number is computed from the date at which the individual joined the CG and the date when he or she retired. Only integer months were calculated and the number of days less than 30 were neglected. For example, an individual who served between 87/09/05 (year, month, day) and 90/02/20, has a calculated value, named MY, of 29 months in active duty. With this information survival functions were developed by paygrade and individual characteristics such as sex, race, marital status and rating. These calculations were performed using the SAS LIFETEST procedure. A program listing is included as Appendix A.

A. SURVIVAL FUNCTIONS

A first step in the analysis of the survival data is the estimation of the distribution of the failure times. The survival distribution function (SDF), also known as the survival function, is used to describe the lifetimes of the enlisted personnel in the CG. The SDF evaluated at $t$, is the probability that an enlisted member will have a lifetime exceeding $t$, that is $S(t) = \text{Prob}(T > t)$ where $S(t)$ denotes the
survival function and $T$ is the lifetime of a randomly selected experimental unit. There are three types of plots commonly used with survival functions: a plot of the estimated survival function against time, a plot of the negative natural log of the estimated survival function against time, and a plot of the natural log of the negative natural log of the estimated survival function against the natural log of time. The last two plots provide empirical checks of the appropriateness of the exponential model and the Weibull model respectively, for the survival data [Ref. 3]. The most important task in the analysis of CG attrition data is the comparison of survival curves. It is of interest to determine whether two or more strata share the same survival functions. The life test procedure can be used with data that may be right censored, (enlisted who actually served in the CG at 30 Sep/90) to compute nonparametric estimates of the survival distributions for each stratum and to perform rank tests for the association of CG attrition with other individual characteristics. The survival estimates are computed within the defined strata levels. Rank statistics and likelihood ratio tests are used to test homogeneity of the survival function over different strata.

In the following section global survival functions are computed for individual characteristics such as sex, marital status, race, paygrade and rating.
1. Enlisted attrition by sex.

As observed from the plot of the survival functions in Figure 3, records for females are limited to a maximum of 180 months since they initially started serving in CG in 1975. Females have demonstrated a higher attrition rate than their counterparts. Comparatively higher attrition rates are found close to 48 months of service for both groups, which corresponds to the end of the first contract. The three tests used for homogeneity indicate that the survival functions for
the two sexes are significantly different. The log survival plot in figure 3, does not present a linear pattern through the origin which is required to support an exponential model, and the plot of the log of the negative log of the survival does not show a linear behavior in support of a Weibull model.

2. Enlisted attrition by marital status.

Figure 4. Enlisted attrition by marital status.

Figure 4 shows the plot of the survival function estimates where the probability to survive in the CG is higher for a married person than for a single member. The
attrition is high in the starting period of service and decreases rapidly for single members once they finish their first four year contract as compared to that of married members. This behavior is inverted when a member reaches 20 years of service. Tests for homogeneity again, indicate that the survival functions for singles and married enlisted are significantly different at alpha $= 0.05$. The log plots support neither exponential nor Wiebull models for the data.

3. **Enlisted attrition by race**

![Enlisted attrition by race](image)

**Figure 5.** Enlisted attrition by race.
The survival functions here show a similar pattern for Blacks and Hispanics. American Indians have the highest attrition. Asian groups show the highest survival probability followed by Caucasians. The plots of the logs conform neither to exponential nor Weibull models.

4. Enlisted attrition by paygrade.

The graph of the survival function estimates in Figure 6, represents the highest attrition on low paygrades until the end of the 4 year contract. The attrition for paygrades E-6
to E-9, increases abruptly after reaching 20 years of service.

The tests for homogeneity indicate that at least one of the survival functions are different from the others. However, in the negative log of the survival function against time plot, Figure 6, the paygrades E-1 to E-3 appear to be candidates for exponential models, while paygrades E-7 to E-9 appear to follow that of a Wiebull model. Note that the general pattern is piecewise linear for each paygrade until the two deep drops around 48 and 240 months.

5. Enlisted attrition by rating

Figure 7 represent the attrition by each rating. There are three of the survival functions that do not follow the pattern of the others. They correspond to the rating 170 (Gunner’s Mate), followed in decreasing order by rating 280 (Telephone Technician) and rating 870 (Health Service Technician). The rating 570 (Aviation Machinist’s Mate) is the group with the highest survival probability followed by rating 100 (Boatwain’s Mate). All survival functions drop sharply around months 48 and 240.

B. SUMMARY

In this chapter, the survival functions were estimated in terms of individual characteristics of the enlisted member. Results obtained are as follow: males and married individuals have higher survival probabilities than their counterparts, respectively; paygrades E-1 to E-5 have higher attrition than
paygrades E-6 to E-9; American Indians have the highest attrition and the Asian members the highest survival functions; rating 170 (Gunner’s Mate) has the highest attrition over all ratings followed by rating 180 (Fire Control Technician); the rating with the highest survival probability is 570 (Aviation Machinist’s Mate); there was a decreasing trend in attrition in the last 4 years of the observation period. The common feature of the survival functions was the significant fall at the end of the four
years service contract and when the enlisted member reach twenty years of service.

In general the survival functions developed in this thesis have similar patterns for the paygrades, marital status and sex to those analyzed in the previous work by LT Blakemore.

For rating and race, this thesis employed more strata than his and the results could not be compared.

It appear that neither Weibull nor the exponential models can be fixed to the survival data.
IV PREDICTION MODEL FOR MONTHLY ATTRITION

This chapter deals with the study of multiple regression models to predict monthly attrition. Monthly attrition is the dependent variable. The potential independent variables for predicting monthly attrition are: monthly attrition in the three previous months, the number of enlisted personnel who joined the CG four years ago and twenty years ago, respectively, seasonality, monthly unemployment rate and time.

A. MULTIPLE REGRESSION MODEL

The reason for using regression models is the desire to forecast. A forecast is a quantitative estimate (or set of estimates) about the likelihood of future events based on past and current information. Two types of forecasting can be applied: the point forecast, that predicts a single number in each forecast period; and the interval forecast, that indicates an interval in which the realized value will lie.

The variables in the following list were used as a set of candidate predictors to forecast monthly attrition (which will be denoted CONTEO).

EXPLANATORY VARIABLES:

YT_1 number of attrition in the previous month
YT_2 number of attrition two months ago
YT_3 number of attrition three months ago
XT_48 number of enlisted who entered four year ago
XT_240 number of enlisted who entered twenty years ago
MM time period starting from Oct/82
TS square term of MM
D1 Dummy variable indicating first semester of each FY (Oct - Mar)
D2 Dummy variable indicating second semester of each FY (Apr - Sep)
RATA monthly unemployment rate of U.S.

XT_48 and XT_240 were included because of the prominences in the survival analysis. These two time periods (48 and 240) showed a significant change in the survival probabilities. The variables MM and TS were used to consider the observational period time trend in the attrition behavior. Two dummy variables, D1 and D2, were included to distinguish the potential difference in attrition behavior in two different semesters. Finally the monthly unemployment rate of U.S. was added to relate the attrition behavior to the condition of U.S. economy. These rates were taken from the Monthly Labor Review, a publication of the US Department of Labor and Bureau of Labor Statistics. Based on these, the initial model considered was as follow:

\[ CONTEO = \beta_1 Y_{T_1} + \beta_2 Y_{T_2} + \beta_3 Y_{T-3} + \beta_4 XT_48 + \beta_5 XT_240 + \beta_6 MM + \beta_7 TS + \beta_8 D1 + \epsilon \]

Note that this model does not contain an intercept term.

Using a backward elimination option of the PROC REG procedure
of the statistical package SAS the following three models were selected to predict the monthly attrition.

**MODEL 1**
\[ \text{CONTEO} = \text{YT}_1 \times \text{XT}_{48} + \text{MM} + \text{TS} + \text{RATA} \]

**MODEL 2**
\[ \text{CONTEO} = \text{YT}_1 \times \text{XT}_{48} + \text{RATA} \]

**MODEL 3**
\[ \text{CONTEO} = \text{YT}_1 \times \text{YT}_2 \]

Details are in Appendix B.

---

**Figure 8.** Validation of model 2 and current CG method.

The predictors variables in the models are significant at level of 5% and the \( R^2 \) for all models is greater than 0.9569.

In order to fit the model, the 96 observations (one for
each month in eight fiscal years from Oct/82 to Sep/90) were used while 33 observations (Oct/90 to Jun/93) were employed to validate the model. As performance criteria, the mean squared error (MSE) and mean relative error (MRE) were used.

The predicted attrition values of each true monthly attrition for FY91, FY92 and the first nine months of FY93 were computed and are in Appendix C. Figure 8 contains the corresponding plots. The mean squared error and the mean relative error of each model are given as follows:

### Mean Squared Error and Mean Relative Error for FY91

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>CG Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MSE</strong></td>
<td>1410.166</td>
<td>515.931</td>
<td>817.225</td>
<td>4331.973</td>
</tr>
<tr>
<td><strong>MRE</strong></td>
<td>0.30325</td>
<td>0.13904</td>
<td>0.18486</td>
<td>1.65557</td>
</tr>
</tbody>
</table>

### Mean Squared Error and Mean Relative Error for FY92

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>CG Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MSE</strong></td>
<td>1286.171</td>
<td>94.482</td>
<td>134.652</td>
<td>1346.880</td>
</tr>
<tr>
<td><strong>MRE</strong></td>
<td>0.38381</td>
<td>0.08839</td>
<td>0.10979</td>
<td>1.01372</td>
</tr>
</tbody>
</table>

### Mean Squared Error and Mean Relative Error for FY93 (9 months)

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>CG Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MSE</strong></td>
<td>2304.833</td>
<td>505.083</td>
<td>533.362</td>
<td>3586.590</td>
</tr>
<tr>
<td><strong>MRE</strong></td>
<td>0.52702</td>
<td>0.27350</td>
<td>0.28616</td>
<td>1.86012</td>
</tr>
</tbody>
</table>
B. RESULTS

The results shown in the previous section indicate that model 2 predicts best the twelve month attrition for FY91, FY92 and the first nine months of FY93 in terms of not only the MSE but also the MRE. In model 2, 96.66% of the total variation in the attrition is explained by the predictor variables such as XT_48, YT_1 and RATA. The correlation among these three predictors are low and the Durbin-Watson statistic is close to two indicating there is no significant first order autocorrelation in the residuals.

In summary, in order to forecast a future monthly attrition one can use:

\[ \text{CONTTEO} = 0.674663 \times YT_1 + 0.170029 \times XT_48 + 6.996215 \times RATA \]

When YT_1 and RATA are unknown at the time of forecast, predicted values of YT_1 and RATA can replace actual YT_1 and RATA.
V. CONCLUSION AND RECOMMENDATIONS

A. CONCLUSIONS

The objective of this thesis was to analyze the attrition behavior and to predict monthly enlisted attrition of the US Coast Guard.

First, survival analysis was used to investigate the attrition behavior of a USCG enlisted in terms of individual characteristics such as sex, marital status, rating, race and paygrade. Results obtained based on the past eight years (Oct/82 to Sep/90) data are as follows: males and married individuals have higher survival probabilities than their counterparts, respectively; paygrades E-1 to E-5 have higher attrition than paygrades E-6 to E-9; American Indians have the highest attrition and the Asian members the highest survival functions; rating 170 (Gunner's Mate) has the highest attrition over all ratings followed by rating 180 (Fire Control Technician); the rating with the highest survival probability is 570 (Aviation Machinist's Mate); there was a decreasing trend in attrition in the last 4 years of observation period. The common feature of the survival functions was the significant fall at the end of the four years service contract and when the enlisted member reach twenty years of service.
Secondly, a regression model was estimated to forecast the monthly attrition: significant predictors are the prior month's attrition, the number of enlistments four years prior and the current unemployment rate as significant predictors. This model explains almost 97% variation in the monthly attrition. Performance of the regression model turns out to be better than that of the current method used in CG.

B. RECOMMENDATIONS
The goals of this thesis were met. Recommendations for further studies in CG enlisted attrition are listed below:

1. Recommendation 1
For future research it will be necessary to have a more manageable enlisted data base in order to facilitate necessary analysis.

2. Recommendation 2
It is recommended to formulate the CG enlisted attrition as a time series model such as moving average models, autoregressive models or a combination of both.

3. Recommendation 3
It is recommended that the forecast model be formulated for each rating.
This source file was used to calculate survival functions for each individual characteristics such as sex, marital status, race, paygrade and rating.

```
OPTION LINESIZE=80;
DATA ONE; SET HELEN.HELEN;
  CENSORED = (SEPAR > 900930 ) ;
  IF PAY = 1 THEN GROUP = 'P1' ;
  ELSE IF PAY = 2 THEN GROUP = 'P2' ;
  ELSE IF PAY = 3 THEN GROUP = 'P3' ;
  ELSE IF PAY = 4 THEN GROUP = 'P4' ;
  ELSE IF PAY = 5 THEN GROUP = 'P5' ;
  ELSE IF PAY = 6 THEN GROUP = 'P6' ;
  ELSE IF PAY = 7 THEN GROUP = 'P7' ;
  ELSE IF PAY = 8 THEN GROUP = 'P8' ;
  ELSE GROUP = 'P9' ;
PROC LIFETEST OUTSURV=OUT1 PLOT=(S) NOTABLE;
  TIME MY*CENSORED(1);
  STRATA GROUP ;
DATA DOS ; SET OUT1 ;
DROP GROUP SDF_LCL SDF_UCL ;
  IF _CENSOR_ = 1 THEN DELETE ;
DATA TRES ; SET DOS(DROP= _CENSOR_ ) ;
PROC PRINT ;
```
APPENDIX B

This is the source file to estimate and validate the 3 regression models.

OPTION LINESIZE=80;

DATA ONE;
  INPUT MM CONTEO XT_48 RATA D1 D2 ;
  CARDS;

;  DATA DOS; SET ONE;
  YT_1 = LAG1(CONTEO) ;
  YT_2 = LAG2(CONTEO) ;
  TS = MM * MM ;
  IF MM < 97 ;
    PROC CORR;
    PROC REG ;
    MODEL CONTEO = YT_1 YT_2 XT_48 MM TS RATA / NOINT DW ;
    PROC REG ;
    MODEL CONTEO = YT_1 XT_48 MM TS RATA / NOINT DW ;
    PROC REG ;
    MODEL CONTEO = YT_1 XT_48 MM RATA / NOINT DW ;
    PROC REG ;
    MODEL CONTEO = YT_1 XT_48 RATA / NOINT DW ;
    PROC REG ;
    MODEL CONTEO = YT_1 YT_2 / NOINT DW ;
DATA NEW ; SET DOS ;
DROP YT_2 ;
IF MM > 108 AND MM < 121 ;
NEWM1 = 0.558604*YT_1 + 0.162241*XT_48 + 9.085655*RATA +
2.094146*MM - 0.025143*TS ;
SQ1 = (CONTEO - NEWM1)**2 ;
RE1 = ABS(CONTEO - NEWM1) / CONTEO ;
NEWM2 = 0.674663*YT_1 + 0.170029*XT_48 + 6.996215*RATA ;
SQ2 = (CONTEO - NEWM2)**2 ;
RE2 = ABS(CONTEO - NEWM2) / CONTEO ;
NEWM4 = 0.785271*YT_1 + 0.193608*YT_2 ;
SQ4 = (CONTEO - NEWM4)**2 ;
RE4 = ABS(CONTEO - NEWM4) / CONTEO ;
DATA NEW1 ; SET NEW ;
DROP MM YT_1 XT_48 TS D1 D2 RATA ;
PROC SUMMARY ; VAR SQ1 SQ2 SQ4 RE1 RE2 RE4 ;
OUTPUT OUT = OUT1 MEAN = MSQ1 MSQ2 MSQ4 MRE1 MRE2 MRE4 ;
PROC PRINT DATA = OUT1 ;
PROC PRINT DATA = NEW1 ;
Correlation Analysis

9 'VAR' Variables: MM CONTEO XT_48 RATA D1 D2 YT_1 YT_2 TS

### Simple Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Sum</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td>96</td>
<td>48.5000</td>
<td>27.8568</td>
<td>4656</td>
<td>1.00000</td>
<td>96.0000</td>
</tr>
<tr>
<td>CONTEO</td>
<td>96</td>
<td>306.2917</td>
<td>106.7245</td>
<td>29404</td>
<td>12.00000</td>
<td>546.0000</td>
</tr>
<tr>
<td>XT_48</td>
<td>96</td>
<td>303.0208</td>
<td>131.3077</td>
<td>29090</td>
<td>38.00000</td>
<td>649.0000</td>
</tr>
<tr>
<td>RATA</td>
<td>96</td>
<td>6.87917</td>
<td>1.56937</td>
<td>660.4000</td>
<td>0.00000</td>
<td>10.90000</td>
</tr>
<tr>
<td>D1</td>
<td>96</td>
<td>0.51042</td>
<td>0.50252</td>
<td>49.00000</td>
<td>0.00000</td>
<td>1.00000</td>
</tr>
<tr>
<td>D2</td>
<td>96</td>
<td>0.51042</td>
<td>0.50252</td>
<td>49.00000</td>
<td>0.00000</td>
<td>1.00000</td>
</tr>
<tr>
<td>YT_1</td>
<td>95</td>
<td>307.90526</td>
<td>106.10694</td>
<td>29251</td>
<td>12.00000</td>
<td>546.0000</td>
</tr>
<tr>
<td>YT_2</td>
<td>94</td>
<td>309.31915</td>
<td>105.77235</td>
<td>29076</td>
<td>12.00000</td>
<td>546.0000</td>
</tr>
<tr>
<td>TS</td>
<td>96</td>
<td>3120</td>
<td>2789</td>
<td>299536</td>
<td>1.00000</td>
<td>9216</td>
</tr>
</tbody>
</table>

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / Number of Observations

<table>
<thead>
<tr>
<th></th>
<th>MM</th>
<th>CONTEO</th>
<th>XT_48</th>
<th>RATA</th>
<th>D1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td>1.00000</td>
<td>-0.61797</td>
<td>-0.41096</td>
<td>-0.92129</td>
<td>-0.10791</td>
</tr>
<tr>
<td>CONTEO</td>
<td>-0.61797</td>
<td>1.00000</td>
<td>0.53645</td>
<td>0.44194</td>
<td>-0.24383</td>
</tr>
<tr>
<td>XT_48</td>
<td>-0.41096</td>
<td>0.53645</td>
<td>1.00000</td>
<td>0.32342</td>
<td>-0.14932</td>
</tr>
<tr>
<td>RATA</td>
<td>-0.92129</td>
<td>0.44194</td>
<td>0.32342</td>
<td>1.00000</td>
<td>0.12041</td>
</tr>
<tr>
<td>D1</td>
<td>-0.10791</td>
<td>-0.24383</td>
<td>-0.14932</td>
<td>0.12041</td>
<td>1.00000</td>
</tr>
<tr>
<td>YT_1</td>
<td>-0.60800</td>
<td>0.81003</td>
<td>0.43099</td>
<td>0.42469</td>
<td>-0.26260</td>
</tr>
<tr>
<td>YT_2</td>
<td>-0.59948</td>
<td>0.69724</td>
<td>0.42722</td>
<td>0.39914</td>
<td>-0.20203</td>
</tr>
<tr>
<td>TS</td>
<td>0.96888</td>
<td>-0.69331</td>
<td>-0.43166</td>
<td>-0.82029</td>
<td>0.09517</td>
</tr>
</tbody>
</table>

33
<table>
<thead>
<tr>
<th></th>
<th>D2</th>
<th>YT_1</th>
<th>YT_2</th>
<th>TS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td></td>
<td>-0.60800</td>
<td>-0.59948</td>
<td>0.96888</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>94</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>CONTEO</td>
<td></td>
<td>0.81003</td>
<td>0.69724</td>
<td>-0.69331</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>94</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>XT_48</td>
<td></td>
<td>0.43099</td>
<td>0.42722</td>
<td>-0.43166</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>94</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>RATA</td>
<td></td>
<td>0.42469</td>
<td>0.39914</td>
<td>-0.82029</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>94</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td></td>
<td>-0.26260</td>
<td>-0.20203</td>
<td>-0.09517</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>0.0101</td>
<td>0.0509</td>
<td>0.3563</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>94</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>YT_1</td>
<td></td>
<td>1.00000</td>
<td>0.80644</td>
<td>-0.68380</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>94</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>YT_2</td>
<td></td>
<td>0.80644</td>
<td>1.00000</td>
<td>-0.67642</td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>94</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td></td>
<td>-0.68380</td>
<td>-0.67642</td>
<td>1.00000</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>94</td>
<td>96</td>
<td></td>
</tr>
</tbody>
</table>
Model: \( \text{MODEL } \) \( \text{CONTEO} = f(\text{XT}_{48}, \text{RATA}, \text{YT}_1, \text{YT}_2, \text{TS}, \text{MM}) \)

NOTE: No intercept in model. R-square is redefined. Dependent Variable: CONTEO

### Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>6</td>
<td>9611402.0244</td>
<td>1601900.3374</td>
<td>485.667</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>88</td>
<td>290254.97562</td>
<td>3298.35200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U Total</td>
<td>94</td>
<td>9901657</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Root MSE 57.43128  
R-square 0.9707  
Dep Mean 306.35106  
Adj R-sq 0.9687  
C.V. 18.74688

### Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | T for H0: Parameter=0 | Prob > |T| |
|----------|----|--------------------|----------------|------------------------|--------|---|
| YT_1     | 1  | 0.569087           | 0.10135072     | 5.615                  | 0.0001 |
| YT_2     | 1  | -0.009225          | 0.10036742     | -0.092                 | 0.9270 |
| XT_48    | 1  | 0.164531           | 0.05108831     | 3.221                  | 0.0018 |
| MM       | 1  | 2.209609           | 0.97305540     | 2.271                  | 0.0256 |
| TS       | 1  | -0.026221          | 0.01081079     | -2.425                 | 0.0173 |
| RATA     | 1  | 8.529324           | 3.38234431     | 2.522                  | 0.0135 |

Durbin-Watson D 1.961  
(For Number of Obs.) 94  
1st Order Autocorrelation 0.009

---

35
Model: \( \text{MODEL CONTEO} = f(XT_{48}, RATA, YT_1, TS, MM) \)

NOTE: No intercept in model. R-square is redefined. Dependent Variable: CONTEO

### Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>9723932.3348</td>
<td>1944786.467</td>
<td>599.485</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>90</td>
<td>291968.66517</td>
<td>3244.09628</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U Total</td>
<td>95</td>
<td>10015901</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Root MSE: 56.95697  
R-square: 0.9708  
Dep Mean: 306.68421  
Adj R-sq: 0.9692  
C.V.: 18.57186

### Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | T for HO: Parameter=0 | Prob > |T| |
|----------|----|-------------------|----------------|-----------------------|--------|---|
| YT_1     | 1  | 0.558604          | 0.07950773     | 7.026                 | 0.0001 |
| XT_48    | 1  | 0.162241          | 0.05018357     | 3.233                 | 0.0017 |
| MM       | 1  | 2.094146          | 0.90905311     | 2.304                 | 0.0235 |
| TS       | 1  | -0.025143         | 0.01015032     | -2.477                | 0.0151 |
| RATA     | 1  | 9.085655          | 3.13547291     | 2.898                 | 0.0047 |

Durbin-Watson D: 1.979  
(For Number of Obs.) 95  
1st Order Autocorrelation: 0.007
Model: \text{MODEL} \quad \text{CONTEO} = f(XT_{48}, \text{RATA}, YT_{1}, \text{MM}) \\
\text{NOTE: No intercept in model. R-square is redefined. Dependent Variable: CONTEO}

\textbf{Analysis of Variance}

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>4</td>
<td>9704026.5996</td>
<td>2426006.6499</td>
<td>707.870</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>91</td>
<td>311874.40041</td>
<td>3427.19121</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U Total</td>
<td>95</td>
<td>10015901</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Root MSE 58.54222  
R-square 0.9689  
Dep Mean 306.68421  
Adj R-sq 0.9675  
C.V. 19.08876

\textbf{Parameter Estimates}

| Variable | DF | Parameter Estimate | Parameter Error | T for H0: Parameter=0 | Prob > |T| |
|----------|----|--------------------|-----------------|-----------------------|--------|
| YT_{1}   | 1  | 0.675515           | 0.06576463      | 10.272                | 0.0001 |
| XT_{48}  | 1  | 0.173063           | 0.05138446      | 3.368                 | 0.0011 |
| MM       | 1  | -0.122828          | 0.16371164      | -0.750                | 0.4550 |
| RATA     | 1  | 7.569406           | 3.16073313      | 2.395                 | 0.0187 |

Durbin-Watson D 2.087  
(For Number of Obs.) 95  
1st Order Autocorrelation -0.050
Model: MODEL CONTEO = f(XT_48, RATA, YT_1)
NOTE: No intercept in model. R-square is redefined. Dependent Variable: CONTEO

### Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3</td>
<td>9702097.409</td>
<td>3234032.4697</td>
<td>948.144</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>92</td>
<td>313803.591</td>
<td>3410.90860</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U Total</td>
<td>95</td>
<td>10015901</td>
<td>58.40288</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Root MSE  | 58.40298 | R-square | 0.9687 |
| Dep Mean  | 306.68421 | Adj R-sq | 0.9676 |
| C.V.      | 19.04336  |          |        |

### Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | T for H0: Parameter=0 | Prob > |T| |
|----------|----|--------------------|----------------|------------------------|--------|---|
| XT_48    | 1  | 0.170029           | 0.05110324    | 3.327                  | 0.0013 |
| RATA     | 1  | 6.996215           | 3.05971911    | 2.287                  | 0.0245 |
| YT_1     | 1  | 0.674663           | 0.06559844    | 10.285                 | 0.0001 |

Durbin-Watson D 2.072
(For Number of Obs.) 95
1st Order Autocorrelation -0.044
Model: \( \text{MODEL} \quad \text{CONTEO} = f(YT_1, YT_2) \)

NOTE: No intercept in model. R-square is redefined. Dependent Variable: CONTEO

**Analysis of Variance**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>9521013.4894</td>
<td>4760506.7447</td>
<td>1150.595</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>92</td>
<td>380643.51056</td>
<td>4137.42946</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U Total</td>
<td>94</td>
<td>9901657</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Root MSE: 64.32285  
R-square: 0.9616  
Dep Mean: 306.35106  
Adj R-sq: 0.9607  
C.V.: 20.99645

**Parameter Estimates**

| Variable | DF | Parameter Estimate | Standard Error | T for H0: Parameter=0 | Prob > |T! |
|----------|----|--------------------|----------------|----------------------|--------|
| YT_1     | 1  | 0.785271           | 0.10156464     | 7.732                | 0.0001 |
| YT_2     | 1  | 0.193608           | 0.10135320     | 1.910                | 0.0592 |

Durbin-Watson D: 2.003  
(For Number of Obs.): 94  
1st Order Autocorrelation: -0.012
APPENDIX C

True and fitted attrition calculated with the selected model and the current method used by USCG.

<table>
<thead>
<tr>
<th></th>
<th>True Attrition</th>
<th>Predicted Attrition Model 2</th>
<th>Predicted Attrition Current Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FY91</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct/90</td>
<td>435</td>
<td>379.454</td>
<td>241.750</td>
</tr>
<tr>
<td>Nov/90</td>
<td>366</td>
<td>310.328</td>
<td>256.250</td>
</tr>
<tr>
<td>Dec/90</td>
<td>251</td>
<td>266.937</td>
<td>367.500</td>
</tr>
<tr>
<td>Jan/91</td>
<td>306</td>
<td>296.300</td>
<td>236.250</td>
</tr>
<tr>
<td>Feb/91</td>
<td>280</td>
<td>302.150</td>
<td>277.750</td>
</tr>
<tr>
<td>Mar/91</td>
<td>342</td>
<td>313.636</td>
<td>303.750</td>
</tr>
<tr>
<td>Apr/91</td>
<td>371</td>
<td>334.110</td>
<td>300.500</td>
</tr>
<tr>
<td>May/91</td>
<td>341</td>
<td>303.537</td>
<td>337.250</td>
</tr>
<tr>
<td>Jun/91</td>
<td>299</td>
<td>310.849</td>
<td>325.750</td>
</tr>
<tr>
<td>Jul/91</td>
<td>383</td>
<td>371.072</td>
<td>342.875</td>
</tr>
<tr>
<td>Aug/91</td>
<td>446</td>
<td>411.705</td>
<td>353.875</td>
</tr>
<tr>
<td>Sep/91</td>
<td>419</td>
<td>370.354</td>
<td>332.375</td>
</tr>
<tr>
<td><strong>FY92</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct/91</td>
<td>391</td>
<td>347.414</td>
<td>262.500</td>
</tr>
<tr>
<td>Nov/91</td>
<td>319</td>
<td>282.724</td>
<td>259.750</td>
</tr>
<tr>
<td>Dec/91</td>
<td>293</td>
<td>280.335</td>
<td>370.500</td>
</tr>
<tr>
<td>Jan/92</td>
<td>338</td>
<td>316.155</td>
<td>273.000</td>
</tr>
<tr>
<td>Feb/92</td>
<td>262</td>
<td>254.189</td>
<td>280.250</td>
</tr>
<tr>
<td>Mar/92</td>
<td>296</td>
<td>286.479</td>
<td>310.625</td>
</tr>
<tr>
<td>Apr/92</td>
<td>303</td>
<td>292.431</td>
<td>304.000</td>
</tr>
<tr>
<td>May/92</td>
<td>283</td>
<td>263.674</td>
<td>340.500</td>
</tr>
<tr>
<td>Jun/92</td>
<td>271</td>
<td>290.604</td>
<td>310.125</td>
</tr>
<tr>
<td>Jul/92</td>
<td>339</td>
<td>340.863</td>
<td>336.125</td>
</tr>
<tr>
<td>Aug/92</td>
<td>342</td>
<td>352.768</td>
<td>348.375</td>
</tr>
<tr>
<td>Sep/92</td>
<td>318</td>
<td>291.120</td>
<td>321.625</td>
</tr>
<tr>
<td><strong>FY93 (nine months)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct/92</td>
<td>292</td>
<td>273.408</td>
<td>307.875</td>
</tr>
<tr>
<td>Nov/92</td>
<td>183</td>
<td>199.190</td>
<td>295.875</td>
</tr>
<tr>
<td>Dec/92</td>
<td>308</td>
<td>284.484</td>
<td>239.750</td>
</tr>
<tr>
<td>Jan/93</td>
<td>237</td>
<td>238.624</td>
<td>267.625</td>
</tr>
<tr>
<td>Feb/93</td>
<td>278</td>
<td>256.763</td>
<td>270.875</td>
</tr>
<tr>
<td>Mar/93</td>
<td>305</td>
<td>291.982</td>
<td>300.875</td>
</tr>
<tr>
<td>Apr/93</td>
<td>244</td>
<td>250.998</td>
<td>296.875</td>
</tr>
<tr>
<td>May/93</td>
<td>355</td>
<td>322.655</td>
<td>327.125</td>
</tr>
<tr>
<td>Jun/93</td>
<td>199</td>
<td>204.822</td>
<td>299.250</td>
</tr>
</tbody>
</table>
LIST OF REFERENCES


<table>
<thead>
<tr>
<th>No.</th>
<th>Distribution List</th>
<th>No. Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Defense Technical Information Center Cameron Station Alexandria VA 22304-6145</td>
<td>2</td>
</tr>
<tr>
<td>2.</td>
<td>Library, Code 052 Naval Postgraduate School Monterey CA 93943-5002</td>
<td>2</td>
</tr>
<tr>
<td>3.</td>
<td>Professor So Young Shon, Code OR/Sh Naval Posgraduate School Monterey CA 93940</td>
<td>2</td>
</tr>
<tr>
<td>4.</td>
<td>Professor Robert R. Read, Code OR/Rd Naval Postgraduate School Monterey CA 93940</td>
<td>2</td>
</tr>
<tr>
<td>5.</td>
<td>USCG Personnel Workforce Planning Office Washington D.C.</td>
<td>2</td>
</tr>
<tr>
<td>6.</td>
<td>LCDR Enrique Oñate Rubiano A.A. 5204 Cartagena Colombia</td>
<td>2</td>
</tr>
</tbody>
</table>