NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

INVESTIGATION OF GOALING MODELS
FOR NAVY RECRUITING

by

John Oliver Donelan

March 1977

Thesis Advisor: D. E. Neil

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): 6 Investigation of Goaling Models for Navy Recruiting.
5. TYPE OF REPORT & PERIOD COVERED: 9 Master's Thesis, ~~March 1977~~
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): 10 John Oliver/Donelan
8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Postgraduate School, Monterey, California 93940
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Postgraduate School, Monterey, California 93940
12. ~~REPORT DATE~~: 11 March 1977
13. NUMBER OF PAGES: 43
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Naval Postgraduate School, Monterey, California 93940 12 43p.
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study reviews previous works which investigated goaling models for the Navy's recruiting effort. Nine models were examined for their goaling potential, and five were selected for further analysis. The results of the analyses supported previous reports that canvassers, unemployment, and the number of qualified military available (QMA) males within a Recruiting District exerted a positive influence on accessions, and that the black QMA

of a district had a negative coefficient associated with it. Because of the high level of significance of the recruiters, it is recommended that a future study attempt to identify the attributes of an effective recruiter.

Approved for public release; distribution unlimited.

INVESTIGATION OF GOALING MODELS FOR NAVY RECRUITING

by

John Oliver Donelan
Lieutenant Commander, United States Naval Reserve
B.S., United States Naval Academy, 1963

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1977

Author: John Oliver Donelan

Approved by: Douglas E Neil, Thesis Advisor

Donald R Bouchoux, Second Reader

Michael G Sovereign, Chairman, Department of Operations Research

J A Schrady, Dean of Information and Policy Sciences

ABSTRACT

This study reviews previous works which investigated goaling models for the Navy's recruiting effort. Nine models were examined for their goaling potential, and five were selected for further analysis. The results of the analyses supported previous reports that canvassers, unemployment, and the number of qualified military available (QMA) males within a Recruiting District exerted a positive influence on accessions, and that the black QMA of a district had a negative coefficient associated with it. Because of the high level of significance of the recruiters, it is recommended that a future study attempt to identify the attributes of an effective recruiter.

TABLE OF CONTENTS

LIST OF TABLES

I. THE PROBLEM

A. INTRODUCTION

Over the years there have been numerous attempts to develop a predictive model that the Navy could utilize to establish recruiting goals for its Recruiting Districts. The purpose of this study is not to come up with "THE" model, but rather to investigate some possible models and to develop those which appear promising.

Specific goals of this project are;

(1) to identify variables which explain more than ninety percent of the variation in the model;

(2) to examine the residuals of these models for possible patterns;

(3) to provide guidelines for future studies and,

(4) to substantiate, when possible, the findings of previous models.

B. PREVIOUS STUDIES

There have been a number of goaling studies completed since the implementation of the All Volunteer Force.

Stephan and Horowitz (1972) and Jehn and Carroll (1973) completed studies directed toward determining the marginal productivity of a Navy recruiter in order to suggest a method of reallocating recruiters to optimize the number of accessions. Lockman (1975) and Neese (1976) oriented their efforts toward goaling models and the variables which best explained the upper mental group accessions.

As a result of the Navy's failure to achieve its recruiting goals from September 1971 to March 1972, Stephan and Horowitz (1972) initiated one of the first of these studies. However, they were hampered by the fact that all the historical data that was available to them was gathered during a time period when the draft was still in effect. Realizing this, their primary objective became one of selecting a model "so that as more reliable data became available the methodology .." could be ".. refined, expanded and validated to be a useful tool in the making of management decisions."

Using the historical data for the number of canvassers and recruits provided specific data as did the seasonally adjusted military eligible population in each district. Other factors such as economic conditions, social attitudes, draft pressures, and recruiting efforts by other services were incorporated into a dummy variable. By their own admission, "such a hodge-podge of effects can have no simple interpretation." Again the realistic goal was to expand the model as these effects became identifiable.

The Cobb-Douglas function, a commonly used aggregate production function (Bowers and Baird, 1971), was deemed an appropriate model to begin with, and marginal productivity was determined by taking anti-logs then partial derivatives. The conclusions were that the total number of accessions was not very sensitive to the distribution of canvassers and that it may be more important "to optimize some of the sub-goals (such as minority recruiting, quality mix or effectiveness) rather than ... total recruits."

Another study which resulted from the same recruiting shortfalls was completed in July 1973 (Jehn and Carroll, 1973). In addition to seeking the cause of the shortfalls, their search also sought to estimate the effects of pay, advertising and recruiters on first term enlistments. While this study realized that it takes a certain length of time for a canvasser to become fully effective, it had several faults. For example, rather than try to identify variables which might have a statistical significance in a regression model, the authors instead chose to use a very similar set of dummy variables and the Cobb-Douglas model which did little to further the work that was done by Stephan and Horowitz (1972). They did recognize, however, that because "of some severe shortcomings in both the data and the regression technique, it would be a mistake to make much of the results" contained in their model by using the Cobb-Douglas function.

Therefore, Jehn and Carroll were unable to make any definitive statements about the effectiveness of increased pay, advertising or recruiters on overall accessions due to the fact that all three were implemented at approximately the same time. They suggested that recruiter selection, training and motivation might be more important than total number of canvassers. The Recruiting Command was implementing a standardized selection and training program at that time, and it is continuing in effect at the present. It was also suggested that management and support personnel be more carefully screened and trained in order to increase productivity. However, this conclusion was not substantiated by analysis, and the number of support personnel proved insignificant in a later study.

Two other suggestions by Jehn and Carroll were to experiment by advertising only in selected areas and to reallocate some recruiters. Elimination of paid advertising invalidated the first suggestion. The second idea has been undertaken by the Recruiting Command, but no data as yet is available to analyze the results.

Continuing chronologically, a report by Bennett and Haber (1974) also centered on allocating recruiters geographically. While no mathematical model was developed, attention was given to a comparison of qualified military available (QMA) as a measure of market size versus market potential. Market potential was described as covering the effects of the density of the area (QMA/square mile), propensity to enlist and the educational attainments of the area. This educational factor was important as a measure of the quality of recruits attainable.

Bennett and Haber divided the country into six geographical areas and calculated the number of enlistees per recruiter. That number was adjusted for those who did not successfully complete basic training, and this figure was defined as net productivity. Reassigning on the basis of net productivity was concluded to be the optimal allocation of recruiters. There was no regression analysis in this report but it did suggest previously unmentioned variables such as propensity and educational levels.

An extensive and in depth study was completed by the Center for Naval Analyses in October 1975 (Lockman, 1975). The specific purpose of the study was to develop a model to explain the differences in enlistment patterns in the Navy Recruiting Districts (NRD). It was the first study to begin to break down the various demographic variables, and it added factors which were subject to the control of the Recruiting Command.

Because of the way in which the NRD boundaries are established, the one geographical unit left intact is the county. Using this unit, the independent variables were identified as black, urban, education, unemployment, mix, per capita income, and net migration into or out of the district. A detailed description of these variables will be given in Chapter 2. The education and the unemployment variables were not available in terms of 17 to 21 year olds, so they were given as median years of education for the district and a percent of the total civilian labor force. Mix was defined as "the proportion of the total civilian labor force employed in manufacturing." The dependent variable was enlistment rates in the districts; i.e. male non prior service enlistments divided by the number of 17 to 21 year olds in the district. All data was based on calendar year 1973.

Lockman (1975) predicted that the urban variable would have a positive coefficient in the regression because more contacts could be made by a recruiter over a specified time period in an urbanized area. Likewise he felt that unemployment would have a positive effect on accessions because of lack of job opportunities. High levels of per capita income were predicted to have a negative effect because of the lower attractiveness of the military compensation compared to civilian opportunities. The expected effects of the other variables were uncertain.

Using regression analysis techniques on linear and logit functions (see Chapter 3 for a description), the results verified his a priori expectations regarding the predicted signs of the three variables. The results of both models were "essentially the same" (Lockman, 1975). The independent variables were said to "explain" 72 percent of the variation in the enlistment rates.

Lockman expanded his model to more narrowly define enlistees as both school eligible and high school graduates. This, he stated, would make the regression "more analogous to a pure supply function since the Navy will accept virtually all these volunteers." In this case the value of R-squared reached .89.

The next step in Lockman's study was to introduce three inputs from the Recruiting Command. These were budget, number of canvassers, and district recruiting goals. Unfortunately, the budget figures were unusable because they represented total expenditures and were not separated into individual items.

The canvasser factor was discussed at length with the emphasis being placed upon which functional form would best estimate the effect of canvassers. Using an argument based on economic theory, Lockman decided that the best functional form should reflect that (1) accessions approach zero as the number of canvassers approaches zero; (2) the canvasser should always have a positive effect on accessions; (3) the size of the positive effect should diminish as the number of canvassers increases and (4) the size of the positive effect will vary depending on the other variables in the district.

Lockman considered linear, logit, quadratic in canvassers, log-interaction, and the Cobb-Douglas functional forms (all of which will be described in Chapter 3) when deciding which would meet the criteria above. The log-interaction met all four, and "happily, in estimation it also performed the best in terms of a higher adjusted R-square, a higher F-ratio, and a lower standard error of estimate." In all of the functional models the canvasser factor had a strongly significant and positive effect on enlistments. The statistical significance of some of the other variables decreased probably due to, in Lockman's words, "collinearity between canvassers and other independent variables."

The final input from the Recruiting Command was that of the district recruiting quotas. It was predicted that this factor would have a negative effect on the model, and this was confirmed by the empirical work. The sample size was ranked by the percent of quota attained, divided in half, and re-examined. The upper half's canvasser factor became statistically insignificant while the lower half's remained positive and significant. This was interpreted to mean that quotas do lessen the effect of canvassers.

Using the log-interaction functional form and all the variables discussed to this point, a regression was performed which yielded an R-squared value of .81. When enlistees were once more narrowly defined as before, this value increased to .91.

Just as the previous studies had done, this one also derived a marginal productivity for recruiters, in this case, 2.02. This value was interpreted as meaning that canvassers had very little significance on enlistments and/or that their method of presenting information to possible recruits was ineffective. The characteristics of the district were concluded to be the primary determinant in enlistments.

Other findings were that the land area and density of an area were insignificant variables in enlistments, and that the black variable was always a highly significant negative factor. After trying to explain this as a distaste for the Navy by blacks, a failure by the Navy to recruit blacks, the Navy having too few black canvassers, and blacks being distrustful of white canvassers, the author concludes by stating, "this ad hoc sociological theorizing does not change the fact that the Navy has not done well recruiting blacks."

Summarizing Lockman's report, he found that canvassers, urban population, educational levels and unemployment were generally factors causing a positive effect on enlistments. Black QMA, per capita income, and quotas had a negative effect. The other variables fluctuated but were generally insignificant.

Commander John Neese of the Navy Recruiting Command completed a multilinear model in April 1976 using fiscal year 75 data (Neese, 1976). In his analysis he used the most comprehensive set of variables to date. His model verified that black QMA was a negative factor in recruiting and that urban population, unemployment and canvasser factor contributed positively. The term "canvasser factor" will be explained in depth in the next chapter.

The findings of the Neese report are what prompted this study. It will be a continuation and expansion into other possible models and will hopefully identify the direction that future studies might take.

II. THE VARIABLES

A. THE NAVY RECRUITING COMMAND

The Navy Recruiting Command is divided into six recruiting areas encompassing forty-three recruiting districts. There are from five to eight Recruiting Districts in each Recruiting Area. A detailed description of the Recruiting Command can be found in Arima (1976).

The Bureau of Naval Personnel determines the number of recruits that the Navy needs each year. In addition to total accessions required, the Bureau also gives the Recruiting Command guidance concerning the maximum number of lower category mental groups permissible and the minimum number of minorities desired.

Each Recruiting Area is assigned responsibility for a percentage of the total recruits required. These percentages are subject to modification during joint meetings with area commanders. The goaling model presently being used by the Recruiting Command (Neese, 1976) is the basis for the initial assignment of percentages.

Lockman (1975) found that quotas were a negative influence on the overall effectiveness of a canvasser, a conclusion which was, he stated, arrived at independently by the Recruiting Command. However, no suitable alternatives to quotas have been decided upon because Congressionally mandated year end-strengths serve as a "penalty" for exceeding recruiting goals. These upper limits pose a serious question as to whether or not a suitable alternative to quotas can be found.

The Research and Analysis Division of the Plans and Policy Department in the Recruiting Command has been accumulating and defining many of the demographic and other variables associated with recruiting for the past few years. Their analytical efforts established realistic goals based on factual data which minimizes alterations to original assignments.

In developing the data dealing with the QMA (qualified military available) in an area, the Recruiting Command updated 1970 census information. Data from the Bureau of Labor Statistics relating to unemployment was studied as was the number of accessions according to achievement scores. Other material examined included geographical, budgeting, advertising and personnel data. The next two sections will discuss the variables rejected and those selected.

B. REJECTED VARIABLES

All data examined was for fiscal year 1975. The FY 76 data was not available when work began on this study because of a change in testing procedures. Half way through the fiscal year the Armed Services Vocational Aptitude Battery (ASVAB) replaced the Armed Forces Qualification Test (AFQT) as the primary test for enlisted screening and classification, and data problems resulted due to a lack of standardization in the reporting process.

Through preliminary analysis and a study of the data sources, a number of variables were rejected for inclusion in this study. Those rejected because of insignificance in the preliminary analyses were (a) the number of support personnel at the recruiting offices, (b) the number of part time offices, (c) the number of vehicle miles driven, and (d) the land area within each recruiting district. That the number of part time offices was insignificant was not surprising considering the fact that twelve of the forty-three districts had no part time offices and another fifteen had three or less. The number of vehicle miles driven was plagued by questionable and unstandardized accountability in addition to being insignificant.

Data problems precluded the use of any budgeting or advertising data. In budgeting, there was little information about the budget breakdown beyond the Recruiting Area level.

Advertising may someday prove significant if it can be broken down into its separate components. For example, if an accurate measure of effectiveness can be developed for public service advertising, magazine ads, or public affairs promotions (the Blue Angels, etc.), accessions may be more accurately predicted. However, it seems unlikely that a true measurement of the impact of these factors is likely in the near future.

Another rejected factor was the number of minority recruiters. Due to the fact that there are relatively few in number in the Navy, their number alone did not provide sufficient data.

A final rejected factor was that of Department of Defense (DOD) share. DOD share was described by Neese (1976) as the number of recruits attained by the other services in the same recruiting area. It was interpreted as an indicator of regional "propensity" although it was realized that it borrowed "much of its significance from the QMA and canvasser factors." It was not used in this study in order to minimize the data which required interpretation.

C. VARIABLES EXAMINED

In general the variables examined here were the same ones that were used in the Recruiting Command's model (Neese, 1976). In that model urban QMA, rural QMA, black QMA, an adjusted unemployment figure, canvasser factor, and DOD share were used. The two major exceptions in this study were that a true unemployment figure was used and, as previously mentioned, the DOD share was omitted. Also included in this endeavor were the number of recruiting stations and the QMA per square mile in each district. These last two factors showed little significance in preliminary work, but they were included here for completeness.

The number of accessions for each district represents only the upper two mental groups of 17 to 21 year old males as defined by their scores on the Armed Forces Qualification Test (AFQT). Mental Group I (MG I) recruits are those who scored in the upper seven percentile in the testing and MG II are those in the 65-92 percentile.

The rural and urban QMA's were estimated based upon a mini sample conducted by the Census Bureau which adjusted the 1970 census figures. The black QMA includes MG III (31-64 percentile) high school graduates since the total of MG I and II represents too small a factor for reliable prediction. The QMA per square mile figure is simply the total of the three QMA's divided by the number of square miles in the district.

The number of Naval Recruiting Stations (NRS) in each district is straightforward. The unemployment figure represents the total percent of people unemployed in the district. An age breakdown for the 17 -21 year old males was not available.

The canvasser factor is a subjective number derived by the Recruiting Command as a measure of a canvasser's usefulness at progressive periods of time. The canvasser is considered 0% effective during his first month of recruiting; 28% effective from one to three months; 70% from four to six months; 90% from seven to twelve months; and 100% effective after a year of recruiting duty. These percentages were taken from the Navy Recruiting Command's FY 75 Program Analyses. Based on first hand information, this breakdown is conservative, but it is the only written estimate of the utility of a canvasser.

Table I presents the data used in this analysis. Accessions are given for MG I and II, QMA in thousands of 17 to 21 year old MG I and II military eligible males, unemployment as a percent, and QMA per square mile in thousands per thousand square miles.

TABLE I

DATA BASE

DISTRICT	ACC	URBAN QMA	RURAL QMA	BLACK QMA	CANV FACT	UNEM	QMA S/M	NRS
Albany	1354	59.4	24.3	5.42	86.7	7.9	3.08	34
Boston	1383	63.9	11.5	2.23	85.2	9.8	1.54	62
Buffalo	1065	68.5	19.4	8.18	80.1	7.9	3.29	49
New York	961	64.1	0.0	16.44	83.3	8.1	26.80	43
Harris.	700	31.7	12.2	4.86	47.4	6.7	2.09	19
Phil.	996	60.2	0.6	8.42	65.4	8.1	11.30	20
Newark	812	45.2	10.1	9.70	59.2	8.0	11.67	32
Montgmy.	548	16.0	11.7	9.34	46.1	5.7	.71	20
Columbia	394	10.1	10.9	10.76	40.2	7.1	.83	17
Jax'vle	690	18.6	9.1	6.35	41.0	6.1	.81	18
Atlanta	499	19.9	12.7	9.78	49.2	7.3	1.11	19
Nashv'le	576	15.9	11.5	7.80	42.5	5.4	.89	14
Raleigh	598	19.5	22.0	12.60	60.4	5.9	1.34	22
Memphis	409	6.6	14.7	11.87	41.0	5.1	.69	18
Miami	849	32.1	3.7	7.06	48.0	7.7	2.01	18
Louisvle	456	17.1	20.4	3.83	60.7	6.2	.73	31
Richmond	425	17.5	13.5	10.98	47.3	4.3	1.17	21
Wash'ton	933	50.0	11.1	19.84	74.8	5.2	4.95	26
Cleve.	880	70.5	15.1	7.88	85.9	7.0	6.35	41
Col'bus	960	59.0	13.2	7.17	79.4	6.3	2.92	30
Pitts.	892	38.4	18.3	6.28	70.3	6.7	2.18	37
Detroit	1635	99.7	21.7	14.33	99.0	11.6	3.09	48
Ind.	617	29.9	20.3	3.86	47.8	5.6	1.98	26
Chicago	1526	115.7	17.3	23.51	131.5	6.2	5.19	59
St Louis	841	41.7	27.7	10.25	74.8	7.6	1.22	33
Des Moi.	668	16.9	34.6	0.91	49.3	4.1	1.04	19
Kan City	732	27.2	25.8	4.48	60.1	5.6	.53	23
Minn.	1265	47.7	39.8	0.81	86.8	6.1	.54	35
Omaha	588	17.0	28.8	0.91	46.9	6.4	.34	26
Milw.	865	35.8	24.6	2.87	55.7	5.9	1.20	31
Denver	914	27.6	20.9	1.55	54.9	4.7	.21	23
Albuq.	602	16.1	16.4	2.49	45.1	7.5	.17	19
Dallas	875	32.3	10.9	5.76	68.4	4.2	.71	33
Houston	494	28.7	4.1	4.35	48.2	4.6	1.45	14
L Rock	423	10.3	12.0	6.42	46.5	6.4	.41	11
New Orl	333	30.6	6.5	10.82	44.8	7.1	.86	12
Okla Cty	490	19.8	15.6	2.11	47.3	4.8	.60	17
San Ant.	598	17.8	9.1	4.25	45.1	5.6	.36	20
Los Ang.	1569	112.6	0.0	9.79	123.1	8.4	9.69	50
Portland	861	30.2	26.6	.55	63.3	7.3	.35	35
San Fran	1873	111.0	28.8	9.52	141.2	8.1	.55	65
Seattle	1405	43.5	27.7	1.35	82.4	8.1	.09	39
S Diego	1225	50.9	6.7	3.97	81.2	9.4	.33	30

III. FORMAT OF THE MODELS

The intent of this chapter is to illustrate the forms of the models examined. It is not intended to discuss their derivation nor the regression techniques used in solving them. For the more curious reader, references are provided so that they may pursue these subjects in greater detail.

All of the models discussed herein were solved using the stepwise regression technique described by the BMD02R section of the Biomedical Computer Program (Dixon, 1973). The stepwise regression procedure was chosen over the backward elimination process, the all possible regressions method and the forward selection process because it is "believed to be the best of the variable selection procedures" (Draper and Smith, 1966).

A. MULTILINEAR MODELS

Multilinear models were chosen as the first to examine primarily because of the work done by the Recruiting Command (Neese, 1976). In that analysis 91% of the variation in accessions was explained by the regression.

The approach used in this study was to investigate a model using all the variables with a floating point intercept, then to expand the model to include the interactive effects between the variables. Next the variables were examined using a zero intercept and again the interactive effects were brought in.

1. <u>No Interaction, Floating Point Intercept</u>

The first model is of the form

$$A = b_0 + b_1X_1 + \dots + b_nX_n + e_i \tag{3.1}$$

where

A = accessions, the dependent variable,

b_0 = the constant or intercept,

X_i = the independent variables listed in Table I, and

e_i = the error term.

An analysis of this form was investigated with all of the variables except QMA per square mile, then another regression was performed including that factor.

2. <u>Interactions, Floating Point Intercept</u>

Interactions are the manner in which two variables combine in their effect on the response or dependent variable. An example of this would be the effect of two different tire treads on wet or dry roads. One may perform better on a dry road and the other may perform better on a wet one. Interaction measures the magnitude of the combined effects of the two variables.

In a model dealing with recruiting, interactive effects possess the capability to explain much of the variability of the dependent variable. Initially all possible combinations of interactions are included with the exception of the QMA per square mile factor. The form of the model is

$$A = b_0 + b_1X_1 + \dots + b_7X_7 + c_{ij}X_iX_j + e_i \quad i \neq j \tag{3.2}$$

The notation is the same as in the first equation with the addition of c_{ij} as the coefficient for the interactive terms. The restriction $i \neq j$ precludes the use of a squared term.

3. No Interaction, Zero Intercept

A zero regression intercept omits b_0 from equation 3.1 which implies that the dependent variable is zero if the independent variables are zero. This is a strong and usually unjustified assumption (Draper and Smith, 1966). However, in this case, it is logical to assume that there would be no accessions if there were no population from which to recruit.

When a zero intercept is used, the variances, covariances, and correlations are computed about the origin rather than the mean (Dixon, 1973). The formula is of the form

$$A = b_1X_1 + b_2X_2 + \ldots + b_nX_n + e_i \qquad (3.3)$$

The notation is the same as in formula 3.1.

4. Interactions, Zero Intercept

By omitting b_0 from formula 3.2, the resulting equation will have interactions and a zero intercept. The resulting form is

$$A = b_1X_1 + \ldots + b_nX_n + c_{ij}X_iX_j + e_i \qquad i \neq j \qquad (3.4)$$

The results of all of the multilinear forms will be discussed in Chapter 4.

B. QUADRATIC MODELS

Lockman (1975) investigated the use of a quadratic term in explaining accessions. In his study only the canvasser term was squared while this study investigates the possibility of the squared term of any factor being significant. There are two forms presented;

$$A = b_0 + b_1 X_1^2 + \ldots + b_6 X_6^2 + e_i \qquad (3.5)$$

and

$$A = b_0 + b_1 X_1^2 + \ldots + b_6 X_6^2 + c_1 X_1 + \ldots + c_6 X_5 + e_i \qquad (3.6)$$

The QMA per square mile term is omitted based upon its weak significance in preliminary analysis. The $c_1, \ldots, c_6$ symbols are the coefficients for the linear terms. As is done in the multilinear models, the results of these analyses will be presented in Chapter 4.

C. MODELS USING LOGARITHMS

Lockman (1975) suggested five models as candidates to derive the marginal productivity of a recruiter. Three of these functions, the logit, the Cobb-Douglas and the log-interaction, explored the possibility of the natural logarithms of the variables explaining the variance in the model.

1. The Logit Function

The logit or logistics function is useful because it has a shape similar to the normal distribution but requires less analytical work. Other uses and properties of this function are described in depth by Johnson and Kotz (1970). In this study the function takes the form

$$\ln A = b_0 + b_1 X_1 + \ldots + b_7 X_7 + e_i \qquad (3.7)$$

2. The Cobb-Douglas Production Function

The form of the Cobb-Douglas production function in this study is

$$\ln A = b_0 + b_1 \ln X_1 + \ldots + b_7 \ln X_7 + e_i \qquad (3.8)$$

This function is one of the most common of the aggregate production functions (Bowers and Baird, 1971).

3. The Log-interaction Model

This model was the only one which met all of Lockman's (1975) requirements (see page 13) for determining marginal productivity. It relates accessions to the natural logarithm of the canvasser factor and the interactions of the other variables with that factor. Its form is

$$A = b_0 + b_1 \ln C + b_2 X_2 \ln C + \ldots + b_6 X_6 \ln C + e_i \qquad (3.9)$$

where

C = canvasser factor.

The results of the models using natural logarithms are discussed in the next chapter.

IV. RESULTS OF THE REGRESSIONS

The results of the stepwise regressions performed in the last chapter are presented here. Only those variables which were significant at a 95% (or better) confidence level are included in the regression equations. The amount of variation explained by the variables in the model is also included.

A. MULTILINEAR MODELS

1. No Interaction, Floating Point Intercept

The equation was

$$A = -29.2 + 5.0\ Ur - 16.1\ Bl + 8.7\ Cn + 33.4\ Un \qquad (4.1)$$

where

A = accessions,

Ur = urban QMA,

Bl = black QMA,

Cn = canvasser factor, and

Un = unemployment.

The order of entry and the increasing R-squared values were canvasser factor (.827), black QMA (.860), unemployment (.890), and urban QMA (.901). This was the model both with and without the QMA per square mile factor included.

2. <u>Interactions, Floating Point Intercept</u>

In this case the equation was

$$A = 55.9 + 9.0\ Cn + .65\ Cn*Un - 13.1\ Bl \qquad (4.2)$$

where the notation is the same as in 4.1 except that Cn*Un is the interaction of the canvasser and unemployment factors. The order of entry was

canvasser factor ($R^2 = .827$),

canvasser-unemployment interaction ($R^2 = .865$), and

black QMA ($R^2 = .892$).

3. <u>No Interaction, Zero Intercept</u>

Using the no interaction, zero intercept model yielded the highest R-squared value of any model. The regression equation was

$$A = 5.4\ Ur - 16.4\ Bl + 8.4\ Cn + 31.6\ Un \qquad (4.3)$$

The factors are the same ones as in equation 4.1, but the order of entry was slightly different. Here the order was

canvasser factor ($R^2 = .971$),

black QMA ($R^2 = .977$),

urban QMA ($R^2 = .980$), and

unemployment ($R^2 = .984$).

4. <u>Interactions, Zero Intercept</u>

The regression equation in this case was

$$A = 10.2\ Cn + .58\ Cn*Un - 12.7\ Bl \qquad (4.4)$$

The variables and their order of entry were identical with the results of the floating point intercept. The amount of variation explained by the model increased from .971 to .978 to .982 as the canvasser, the interaction of the canvasser and unemployment, and the black QMA factors entered the equation.

B. QUADRATIC MODELS

Using only the squared terms gave

$$A = 223.2 + .066\ Cn^2 + 5.25\ Un^2 + .17\ Ru^2 \qquad (4.5)$$

where the previously undefined characters are

Cn^2 = the canvasser factor squared,

Un^2 = the unemployment factor squared, and

Ru^2 = the rural QMA factor squared.

The entry order was the same as appears in the equation, and the R-squared values increased from .765 to .833 to .855.

Using both the squared and linear terms together in a stepwise regression model produced

$$A = 129.3 + 8.5\ Cn + 3.2\ Un^2 - 32.\text{[illegible]}\ 4.5\ Ur + .87\ El^2 \qquad (4.6)$$

with the only previously undefined variable being El^2, the black QMA factor squared. The variables in equation 4.6 are listed in the order that they entered the model. The R-squared values were .827, .866, .896, .905, and .911.

C. MODELS USING LOGARITHMS

1. The Logit Function

The significant variables in this regression were

$$\ln A = 6.1374 + .0146\ Ur - .0217\ Bl + .0057\ Ru \qquad (4.7)$$

where

$\ln A$ = the natural logarithm of accessions.

The intercept value in this equation is the equivalent to 462.8 accessions. The R-squared values were, in ascending order of the entry of the variables, .743 to .820 to .833. This model had the lowest R-squared value of any of the models examined.

2. The Cobb-Douglas Production Function

The Cobb-Douglas function took the following form;

$$\ln A = 3.649 + .396 \ln Ur - .092 \ln Bl + .434 \ln Cn \qquad (4.8)$$

where

ln = the natural logarithm of the variables.

The variables are listed in the order that they entered the regression. The amount of the variation of the model which was explained by these variables was .834, .865, and .886.

3. The Log-interaction Model

Of the models using logarithms, this function was the only one achieving an R-squared value in excess of .90. The equation resulting from the regression was

$$A = -1382.4 + 455.3 \ln Cn + 7.5\, Un*\ln Cn - 3.0\, Ll*\ln Cn + 1.5\, Ur*\ln Cn - 1.8\, Q*\ln C \quad (4.9)$$

where Q*ln C is the interactive effect between QMA per square mile and the natural logarithm of the canvasser factor. The variables are listed in the order of their entry, and the R-squared values started at .827 and increased to a final value of .905.

V. ANALYSIS OF THE RESULTS

Results of the analyses of the various models developed in the previous chapter showed that the canvasser factor was the most important variable in explaining accessions. Other positive contributions were the unemployment rate and the urban QMA. The black QMA had a negative influence on the number of MG I and II recruits.

Five of the models examined had R-squared values which exceeded .90. The five were (1) linear with a floating point intercept and no interactions; (2) linear with a zero intercept and no interactions; (3) linear with a zero intercept and interactions; (4) quadratic and linear terms combined and (5) the log-interaction model. They are referred to as cases 1 through 5 for the remainder of this chapter.

A. GENERAL

Three of the variables being examined appeared, in one form or another, in all five models. They were the canvasser factor, the black QMA factor, and the unemployment figure.

By far the most significant variable was the canvasser factor. It was the first variable to enter every model, and it "explained away" a minimum of 82.7% of the variation in the model. It had a positive coefficient in all cases, and therefore it agreed with the findings of Lockman (1975) and

Neese (1976). The consistency of the results suggests that the canvasser factor be examined in greater detail in an attempt to break it down into its separate parts.

In an effort to identify the variables which interacted most significantly with the canvasser factor, a linear regression was performed using only the interactive effect of the canvasser and demographic variables. The resulting equation was

$$A = 333 + .786\ Cn{*}Un + .064\ Cn{*}Ru + .048\ Cn{*}Ur - .135\ Cn{*}Bl.$$

R-squared equalled .879 and the order of entry was as shown in the equation. A possible interpretation of this regression is that it is a relative indication of propensity in a district. In other words, high unemployment may influence more people to join the military than any other factor, and people living in rural areas may be more inclined to join the Navy than those in urban areas.

The black QMA factor was the second variable to enter in cases 1 and 2, and the third to enter the stepwise regression of cases 3, 4 and 5. In case 5, the log-interaction model, the interactive effect of the black QMA and the natural logarithm of the canvasser factor was the actual variable which entered. The negative coefficient associated with the black QMA variable in all models was also in agreement with the Lockman (1975) and Neese (1976) studies.

In recent years the minority goals established by the Bureau of Naval Personnel have acted as a negative influence on accessions; i. e., the higher the minority goal, the lower the overall number of MG I and II recruits. The negative coefficient of the black QMA suggests that this is indeed the case. This is in direct conflict with the findings of a youth attitude survey conducted for the

Defense Department in 1972 (Nadel, 1973) which indicated that a higher percentage of black youths were willing to join the military than white. Unfortunately, however, this report did not break down the percentages by service.

The final variable which appeared in all of the models was the unemployment figure. In both of the linear models which did not examine interactive effects, it was the third or fourth variable to enter the equation. Its interactive effect with the canvasser factor or the natural logarithm of that variable caused it to be the second factor to enter in cases 3 and 5. The interactive effect of unemployment and the canvasser factor infers that more concentrated efforts by recruiters in areas of high unemployment might yield substantial dividends. In the quadratic model it was one of two variables whose squared term entered the regression, and it was the second factor to enter. The positive coefficients of the unemployment variable supported the findings of Lockman (1975) and Neese (1976).

In cases 1, 2 and 4, the urban QMA was a significant variable, and the interaction of the logarithm of the canvasser factor and the urban QMA was meaningful in the log-interaction model. Positive coefficients were associated with this factor which once again was in agreement with previous studies (Lockman, 1975 and Neese, 1976).

The equations in the first four models had intercept values that were of very little consequence in the computed accession figures relative to the actual range, 333 to 1873. Interpretation of the -1382.4 intercept in case 5 is unclear although of more interest in regressions are the loci of the curves which define the 95% confidence limit of the range of accessions (Draper and Smith, 1966). It should be pointed out, though, that the computed accession values for that model were from 348 to 1842, and that range compared favorably with the actual figures.

B. RESULTS OF RESIDUAL TESTING

In this section the results of residual testing are reported for the five cases examined. The first part discusses the results of testing the residuals for normality, and part two reports the Durbin-Watson test results for serial correlation. The final part discusses the Recruiting Districts whose residuals differed by more than two standard deviations from their expected values and possible reasons for the differences.

1. Test For Normality

In testing the residuals for normality, an unbiased mean square error was computed for all models. A 95% confidence level was then determined for each case, and the number of districts whose residual values were outside this figure were counted. Table II presents these findings.

Table II: Residual Test For Normality

Case	M.S.E.	Range	No. Outside Range
1	15,824.3	±246.6	3
2	15,440.1	±243.6	3
3	16,680.4	±253.1	1
4	14,475.4	±235.8	2
5	15,480.7	±243.9	2

With 43 data points, a 95% confidence interval should include approximately 41 of these districts. Therefore, the results imply that the residual values are distributed normally.

2. The Durbin-Watson Test

The Durbin-Watson (D-W) test hypothesizes that the correlation coefficient equals zero against the alternative that it is greater than zero (Wonnacott and Wonnacott, 1972). The procedure followed was to compute a value for the D-W statistic then to extract an upper limit value (d_u) from a table giving serial correlation values (Pindyck and Rubinfield, 1976). The test was $2 < D\text{-}W < 4 - d_u$. Since all cases met the criteria at the 95% confidence level, the null hypothesis that no serial correlation was present was accepted. Table III summarizes the results.

Table III: Durbin-Watson Statistics

Case	D-W Statistic	d_u
1	2.21	1.63
2	2.18	1.63
3	2.28	1.58
4	2.19	1.69
5	2.18	1.69

3. Districts With Extreme Deviations

Several districts had residual values which were two standard deviations above or below their actual level of accessions. The Louisville District was below their computed value in all five cases, and the Cleveland area was below theirs in all models except case 3. The Seattle District exceeded its computed accessions by more than two standard deviations in the linear model with a zero intercept.

The interpretation of this information is unclear. It may be an indication of poor or superior performance, or it may be due to data errors resulting from incorrect reporting procedures. Another possibility is that there may be a significant explanatory variable which is unique to those districts but is not included in the regression models.

Investigation of the Recruiting Districts which deviated substantially from the predicted values may provide useful information in future models. However, the data used in this study is obsolete for this purpose. A better use for the regression models described here is to establish goals for future recruiting.

VI. RECOMMENDATIONS

Unfortunately, there are a number of uncontrollable variables that effect the Navy's recruiting effort. Policy changes within the Navy and the Department of Defense require continuous revision of goaling models. For example, if the Bureau of Naval Personnel increases the number of MG IV recruits allowed to enlist, the canvasser's job becomes simplified because there have always been more applicants of this category than were permitted to join.

On the other hand, an increase in the minority goal has usually had the opposite effect, apparently because it has detracted from the effort to recruit the more easily accessible white MG I and II applicants. The results of the regressions performed in this study support that theory.

Personnel transfers of experienced recruiters back to sea duty also has an adverse effect on recruiting. Longer assignments to recruiting billets could partially compensate for this.

Another problem area in developing a goaling model is the collection of reliable data. The Research and Analysis Division of the Recruiting Command has done much to standardize reporting procedures, but there are a few districts whose data remains highly suspect. Re-emphasizing the importance of accurate reporting to those districts could eliminate this problem.

The models developed in this study pointed out that the recruiter is by far the most important variable in recruiting. A need exists, however, to define the characteristics that identify a "good" recruiter.

A possible starting point would be to administer personality tests to all recruiters on board. Personality traits of the more productive canvassers can be compared to those whose performance has been substandard to attempt to identify significant differences. If any are detected, then they can be used as an initial set of criteria in selecting future recruiters. There are presently over 300 commercially available personality tests from which to choose (Salvendy and Seymour, 1973).

A second use of the results would be to identify the traits of the recruiters who were most successful at recruiting minorities. Information of this nature could be used to reduce or eliminate the negative effect of minority recruiting. Since the majority of personnel oriented tests have been developed for white, middle class people, caution should be used in the selection and interpretation of the tests in order to minimize cultural biases.

Analysis of information concerning the age, marital status, and educational backgrounds of the recruiters may also prove useful. All of these recommendations are intended solely to provide possible starting points for future studies in this area.

Experimentally re-allocating recruiters should be continued with the results being carefully monitored. Trends in this data may provide insight into optimal allocation. Concentrating the limited advertising assets available into specific districts could also provide useful information.

A final recommendation is that this type study be expanded to include all accessions. Updating of all QMA data to include the MG IV category would be necessary, but such a study could provide a comparison between "quantity" and "quality" (MG I and II) accessions in each district.

As was stated at the onset, this study was not designed to provide "THE" model for Navy goaling. Rather it gives a comprehensive overview of what has been done to date, and it provides regression models which may be used as the reader deems appropriate. Because of the ever changing data, policies, and canvassers associated with Navy recruiting, annual updating of goaling models will always be necessary.

LIST OF REFERENCES

1. Arima, J. K., _A Systems Analysis of Navy Recruiting_, Navy Personnel Research and Development Center Special Report 76-9, San Diego, California, 1976.

2. Bennett, J. T. and Haber, S. E., "The Allocation of Recruiters Among Spatial Areas," _Naval Research Logistics Quarterly_, p. 649-658, December 1974.

3. Bowers, D. A. and Baird, R. N., _Elementary Mathematical Macroeconomics_, Prentice-Hall, Incorporated, Englewood Cliffs, New Jersey, 1971.

4. Dixon, W. J. (editor), _Biomedical Computer Programs_, University of California Press, Berkeley, California, 1973.

5. Draper, N. R. and Smith, H., _Applied Regression Analysis_, John Wiley and Sons, Incorporated, New York, New York, 1966.

6. Jehn, C. and Carroll, H. E., _Navy Recruiting in an All-Volunteer Environment_, Center for Naval Analyses Research Contribution 235, Arlington, Virginia, 1974.

7. Johnson, N. L. and Kotz, S., _Continuous Univariate Distributions-2_, Houghton Mifflin Company, Boston, Massachusetts, 1970.

8. Lockman, R. F., _Productivity of Navy Recruiting Districts_, Center for Naval Analyses Paper (CNA) 1526-75, Arlington, Virginia, 1975.

9. Nadel, A. B., Attitudes of Youth Toward Military Service, Manpower Development Research Program Report MR 73-4, Alexandria, Virginia, 1973.

10. Neese, J. F., Goaling Study, paper presented at Chief, Navy Recruiting Command's office, Arlington, Virginia, 19 April 1976.

11. Pindyck, R. S. and Rubinfield, D. L., Econometric Models and Economic Forecasts, McGraw-Hill, Incorporated, New York, New York, 1976.

12. Salvendy, G. and Seymour, W. D., Prediction and Development of Industrial Work Performance, John Wiley and Sons, Incorporated, New York, New York, 1973.

13. Stephan, R. A. and Horowitz, S. A., Recruiting Quotas and Distribution of Canvassers, Center for Naval Analyses Paper (INS)526-72, Arlington, Virginia, 1972.

14. Wonnacott, T. H. and Wonnacott, R. J., Introductory Statistics for Business and Economics, John Wiley and Sons, Incorporated, New York, New York, 1972.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Navy Recruiting Command, Code 22 4015 Wilson Boulevard Arlington, Virginia 22203	2
3.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
4.	Department Chairman, Code 55 Department of Operations Research Naval Postgraduate School Monterey, California 93940	2
5.	Asst. Professor D. E. Neil, Code 55 Ni Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
6.	Lcdr. J. O. Donelan, USNR-TAR 15 Revere Road Monterey, California 93940	1