AD-A070 451 GENERAL RESEARCH CORP MCLEAN VA OPERATIONS ANALYSIS GROUP F/G 5/9
DESIGN AND DEVELOPMENT OF ELIM-IV/COMPLIP-G3.(U)
DEC 78 P B MCWHITE, K D MIDLAM, A K BOCAST MDA903-77-C-0366

UNCLASSIFIED CR-239 NL

1 OF 2
AD
A070451

LEVEL

CR-239

Design and Development of ELIM-IV/COMPLIP-G3

Final Report

By:

P.B. McWhite
K.D. Midlam
A.K. Bocast
J.W. Allen
S.G. Ross
P.J. Neilsen
G.O. Swanson
W.A. Adams
H. Weigel

Contract No. MDA 903-77-C-0366

31 December 1978

OPERATIONS ANALYSIS GROUP
GENERAL RESEARCH CORPORATION
A SUBSIDIARY OF FLOW GENERAL INC.
7655 Old Springhouse Road, McLean, Virginia 22102

Prepared For:

Office of the Assistant Secretary of the Army (MRA&L)
The Pentagon
Washington, D.C. 20310

CR-239

2

Design and Development of ELIM-IV/COMPLIP-G3

Final Report

By:

P.B. McWhite
K.D. Midlam
A.K. Bocast
J.W. Allen
S.G. Ross
P.J. Neilsen
G.O. Swanson
W.A. Adams
H. Weigel

Contract No. MDA 903-77-C-0366

31 December 1978

OPERATIONS ANALYSIS GROUP

GENERAL RESEARCH CORPORATION

A SUBSIDIARY OF FLOW GENERAL INC.

7655 Old Springhouse Road, McLean, Virginia 22102

Prepared For:

Office of the Assistant Secretary of the Army (MRA&L)
The Pentagon
Washington, D.C. 20310

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (14) CR-239

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle) (6) Design and Development of ELIM-IV/ COMPLIP-G3.

5. TYPE OF REPORT & PERIOD COVERED (9) Final rept. 1 Oct 77-30 Sept 78.

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s) (10) P. B./McWhite, K. D./Midlam, A. K./Bocast, J. W./Allen, S. G./Ross. [redacted]

8. CONTRACT OR GRANT NUMBER(s) (15) MDA903-77-C-0366

9. PERFORMING ORGANIZATION NAME AND ADDRESS
General Research Corporation
7655 Old Springhouse Road
McLean, Virginia 22101

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of the Ass't Sec'y of the Army
(MRA&L), The Pentagon
Washington, D. C.

12. REPORT DATE (11) 31 Dec[redacted]78

13. NUMBER OF PAGES 116

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 120p.

15. SECURITY CLASS. (of this report) Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Unlimited

DISTRIBUTION STATEMENT A
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)
Army manpower, strength forecasting, linear programming, enlisted attrition, gains forecasting

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
This study provided several enhancements and extensions to the ELIM-COMPLIP manpower strength forecasting system. Primary enhancements included expanding the maximum number of characteristics groups from 4 to 40 and the development of a linear programming methodology for Career Force Optimization over a fixed years of service distribution objective.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

411 050

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

CONTENTS

CONTENTS

PRECEDING PAGE BLANK-NOT FILMED

INTRODUCTION

Under contract MDA903-77-C-0366, General Research Corporation (GRC) has accomplished a series of tasks to enhance and extend the capabilities of the ELIM-COMPLIP manpower programming system. The resulting enhanced system has been designated as ELIM-IV/COMPLIP-G3.

This report reviews the objectives of this study and the technical foundations of the efforts undertaken by GRC to accomplish the tasks required to satisfy these objectives.

OBJECTIVES OF THE TASKS

Task 1 - Second-Timer Aging

Expand the Inventory Projection Module to include a two-dimensional array for the First-Term Second Timers so that this population is segmented by months to ETS and months of service. Modify the Inventory Projection Module so that loss rates which are generated by the system are properly applied to these two-dimensional arrays. Modify the Data Processor and Factor Development Modules so that loss rates applicable to elements of these two-dimensional population arrays are properly generated, similar to those currently generated for the remaining population arrays.

Task 2 - C-Group Expansion

Expand the current number of characteristic groups for non-prior service (NPS) gain projections from 4 to a possible maximum of 40. The increased number of characteristic groups would allow the user to group on enlistment terms in addition to groupings on sex, education, and mental category. The period for which loss data relevant to these characteristics are developed and retained is to be expanded from 21 to 55 months. Each of the characteristic groups is to include the following loss categories: Expeditious Discharge, Trainee Discharge, Other Adverse, Physical Disqualification, Research, Other, Dropped from Military Control, Immediate Reenlistment, ETS, and Extension. Information on marital status, number of dependents, and other variables to be selected by the user will be stored on the new data base for retrieval purposes. The user will also be provided the capability to vary the number of characteristic groups used in a forecast.

Task 3 - User-Defined Quality Classes

Provide program specifications to USAMSSA which would allow a simplified user selection of the data from the large non-prior service cohort file when the large numeric data base is being developed. USAMSSA will implement the specifications and bring the programs on line. The result will be the capability to create the large numeric data base from scratch by creating an all-months version of file MT63A for input into Program QP8.

Task 4 - Non-Prior Army Service Gains

Provide the capability to automatically pass Non-Prior Army Service Gains data via throughput files to the Report Generator Module (RGM) and allow the user to directly input this gain category projection into the Inventory Projection Module (IPM).

Task 5 - Career Force Optimization

Develop the capability to optimize the various types of enlisted gains to the career population category in order to achieve and maintain a designated career strength. In addition, develop the capability to optimize simultaneously the career strength by years of service in order to achieve a user-specified years-of-service profile. This optimization should take place within the linear programming portion of the ELIM-COMPLIP model and allow a wide range of analyst influence. During optimization the appropriate internally generated loss rates must be applied to the various population elements. The years-of-service array should be broken out individually (by each year) for at least the first 20 years of service.

In order to optimize the YOS distribution an objective function will be added to COMPLIP which minimizes the weighted sum of the deviations of the actual from the planned career force strength, by YOS, subject to user-specified constraints on total strength, the supply of NPS enlistees, PS enlistees by YOS, and the first-term reenlistment gains by years of service.

When the user selects the career force optimization mode, COMPLIP will determine the first-term reenlistments which minimize the above objective function. Then, new reenlistment factors will be computed from the optimal reenlistment gains. A second pass through the IPM will apportion the optimum reenlistment gains by C-group and compute careerist losses.

Task 6 - First-Term Tracking

Develop the capability to track and observe all enlisted personnel, including extendees, through their initial enlistment until they reenlist or separate. This is to provide a more accurate accounting of First and Career Reenlistments. In the existing ELIM-COMPLIP system, when First Termers with enlistments of 4 to 6 years reenlist after 36 months of service they are counted as Career Reenlistments as opposed to First Reenlistments.

Task 7 - First Term to Careerist Transitions

Develop the capability to display by month and year the dynamic flow of personnel as they age from the First Term to the Career population category. The objective of this and of Task 6 is to provide the analyst with the information needed to analyze, project, and control the career strength.

Task 8 - Loss Rate Projection Methods

Develop the capability to use where applicable a least squares technique for fitting an exponential equation to historical data and then use that equation for loss projection. The Factor Development module should be modified and provided with extensive but simple-to-apply user controls on the least squares coefficients.

In addition, other statistical techniques as well as those used in the NPS Gains Module are to be examined for possible inclusion in

ELIM. The analysis should focus on developing seasonality constraints for various types of losses and specifically for non-disability retirement.

Currently, the loss rates are projected in FAC by direct extension of the last point in an exponentially smoothed history of loss rates. Extensive capabilities are provided for user modifications of these loss rates. In this task, GRC is to explore alternative methods of automatic generation of loss rate projections. Of special interest are least squares fits of exponential forms to the historical data, and Bayesian methods of adaptive forecasting. Methods other than the exponential least squares technique, found to be useful alternatives to the current method, will be recommended for future incorporation into the FDM as additional options under user control.

Task 9 - Loss Factor Movement

Modify the Factor Development Module to permit the movement of loss factors between categories of enlisted personnel. This capability will allow the analyst to select existing factors within the loss matrix that apply to one category of enlisted personnel and easily move them to another portion of the matrix to serve as new loss factors for a new or existing personnel category.

Task 10 - Qualitative Data Processor

Develop a qualitative data processor (QDP) necessary to implement Task 2, C-Group expansion, in the ELIM-COMPLIP system. Additional data items on an individual's record will be maintained and updated through use of the Enlisted Master File for management information purposes and future characteristic group breakouts. The data will reside on a large tracking cohort file beginning with data from January 1972. New data elements will include marriage/dependents, changes in marriage/dependents, current MOS, and enlistment expiration time. The QDP will be developed to simplify production requirements and minimize and simplify user controls.

Task 11 - Key Word Processor

Develop a key word processor to create the 40 characteristic groups for input to the ELIM-COMPLIP system from the cohort file developed in Task 10. It should be based upon logical selection criteria using an English language type command structure to simplify user control and TSO operations.

Task 12 - Cohort Targeting Routine

Examine the feasibility and acceptability of a gains cohort loss analysis and provide ELIM with the capability of tracing loss rates for yearly cohorts and targeting for their loss projections. The result of this type of analysis should be more accurate loss projections and the ability to distinguish yearly gain cohorts and their related losses from one another. The user should have the capability to place constraints on total losses for any given yearly cohort group. Develop graphs of the loss patterns of the individual yearly cohorts as well as summary graphs.

In this task, GRC is to develop a system to permit the user to modify projected loss rates for first timers by accession cohort. The system is to operate independently of the ELIM-COMPLIP production flow. It will accept as input the standard loss rate projections by C-group by month of service; construct accession cohort loss curves by month of service; accept user modifications to those curves; and, when ordered, revise the standard projected loss rates in accordance with the results of the users' modifications of the cohort loss curves. The system will operate either interactively or in a batch mode.

Task 13 - NPS Gains Forecasting

Modify the existing NPS Gains Module to provide the data necessary to use the expanded number of characteristic groups provided by Task 2. This requires the capability to provide NPS gain forecasts for the expanded number of characteristic groups as well as projections of seasonality patterns and availability constraints of supply-limited groups.

The NPS Gains Module was created as part of the ELIM-III/COMPLIP-G2 development effort. It provides COMPLIP with one-year-ahead monthly forecasts of accessions by C-group using a historical data base derived from the monthly transaction tapes and a modified least squares time-series methodology with nonlinear deseasonalizing. It is capable of analyzing, graphing, and projecting gains by month for any of the 150 accession groups in the Large Numeric Data Base. Since only four C-groups can be handled in ELIM-III/COMPLIP-G2, the NPS Gains Module always requires that the groups analyzed and projected be aggregated up to the maximum of four C-groups defined in the Small Numeric Data Base.

With the expansion of the C-groups and the new data base planned for current development, changes will be required in the NPS Gains Module. For the most part these changes will be only those needed because of the C-group and data base changes. A few small enhancements will, however, also be accomplished. The one-year-ahead restriction will be relaxed so that projections can cover the full COMPLIP projection timeframe. Projected gains by C-group will be available to the IPM as well as to COMPLIP. One or more new subroutines will be provided to permit user modifications of the projections similar to those available for modifying loss rate projections.

Task 14 - Matrix Generator Enhancements

Develop the capability within the matrix generator to express linkages between the various characteristic groups as ratios or percentages of each other (e.g., Mental-Category-IV High-School Diploma-Graduate equals 10% of all High-School-Diploma-Graduate Accessions during a particular projection period). Also develop the capability to equate one of the characteristic groups to WACs. Accommodate the special constraints of the budget year by developing different capabilities for the budget year and the remaining years.

The existing ELIM-COMPLIP system is to be modified by GRC to provide the user with the capability within the matrix generator to specify the accession level of first-term enlistments of one characteristic group as a fraction of the accession level of another characteristic group and with the capability to equate one of the characteristic groups to WACs with user-supplied accession levels.

Task 15 - Support and Training

Provide assistance in adapting all related computer programs to the designated government computer and assist in training users.

IMPACT OF THE TASKS ON THE SYSTEM

Many of these tasks address needs of particular modules of the system and have little, if any, impact on the other parts of the system. Technical discussions of the accomplishments of such tasks are provided in separate appendixes to this report. On the other hand, some tasks required significant modifications to most, if not all, parts of the system. These tasks include:

- Task 1: Second-Timer Aging
- Task 2: C-Group Expansion
- Task 6: First-Term Tracking
- Task 7: First-Term to Careerist Transitions

The changes in system design and operation required to accomplish these tasks are discussed in the following.

The changes in the inventory projection methodology and data arrays described here enable ELIM-IV to satisfy the requirements of several of the tasks described above. All personnel, including extendees, are tracked through their entire initial enlistments until they either reenlist or separate. The number of C-groups is expanded so that term of service can be used as a classification characteristic. Provisions have been made for operating on up to 40 C-groups. Personnel are projected by C-groups through the first 55 months of service.

The inventory projection procedures can be divided into those used for projecting populations--whose strength levels are not determined by COMPLIP, and those used for projecting rates--for segments of the inventory whose strength levels are determined by COMPLIP.

Projection of Populations

To handle the projection of many more C-groups for many more months, the FTI array has been expanded and the Second-Timer (STI) and careerist arrays have been redefined. The ELIM-II/III array used for 4 C-groups and 21 months of service is replaced. The three area matrices (CONUS, Short-Tour, and Long-Tour) and the AUS matrix are unchanged.

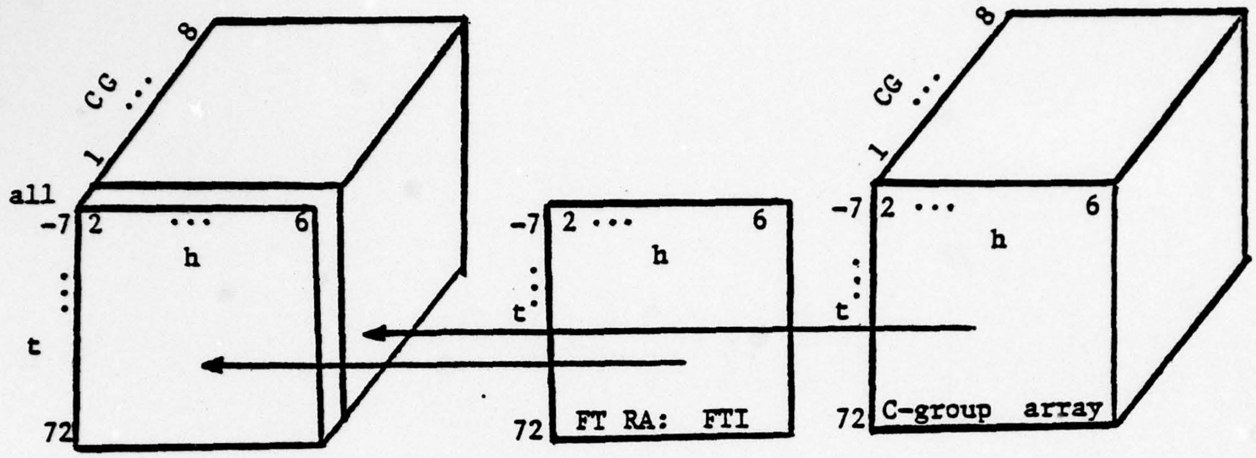
Changes to FTI Array

The most notable change is in the C-group arrays. These arrays in ELIM-II/III are two-dimensional, by C-group (CG = 1,2,3,4) and month of service (m = 1,2..20,21). A third dimension has been added to this array to accommodate the term of enlistment. The summary numbers across C-groups are to be retained in the non-qualitative First-Timer (FTI) arrays. These are not merely the summation of the numbers in the C-groups but are rather the total counts used to normalize C-group values.[1] That is, we are not merely spreading the data in the current array over an additional dimension. Figure 1 illustrates this concept.

To minimize the increase in the size of the IPM, the maximum number of C-groups for each term of enlistment is limited to sixteen each for terms 3 and 4, and four each for terms 2, 5, and 6 (subject, in total, to the 40 group overall limitation). Over the short term, it is planned that a common set of eight C-groups will be defined for terms 3 and 4, and that terms 2, 5, and 6 would have one C-group each.

---

[1]The totals in the FTI array do not equal the sums of the C-groups, because some observations are dropped during creation of the C-group array.



CG = Characteristic Group; h = month of service;
t = months to ETS; FT = First Term; RA = Regular Army; FTI = First Timers

Figure 1. Relationships Between C-Group and FTI Arrays

Changes to STI Array

During historical tracking and inventory projection, populations remain in the FTI and C-group arrays until they either extend, reenlist, separate, or, in the case of C-group arrays, have served 55 months. The STI array is redefined as an extendee-only array. It holds all extendees until they either reenlist or separate or have served 55 months. Administrative gains, reservists who enter active duty, prior service gains who have not completed a first reenlistment, and other miscellaneous gains are added to the revised "Reenlisted" array by year of service. Enlistees in the STI array are classified by year of service and months to ETS.

Changes to Careerist Array

The present careerist array has been redefined as a reenlisted array. Any first-timer or extendee who reenlists will move to the reenlisted array. Since reenlistment can occur before the 36th month of service, the reenlisted array has been expanded to include people with fewer than three years of service. The array has been expanded to present data by individual years of service for non-retirement eligible personnel; that is, the length of service dimension is increased from two classes (NRE, RE) to as many as 21 classes (1,2,...20, 21+). The time-to-ETS dimension will not be changed. The reenlisted array is intended to contain every person who is not either a First-Timer or a First-Timer on an extension.

Projection of Rates

The IPM/QIPM applies loss rates to retention rates as well as to populations. Retention rates are used to calculate NPS gains in COMPLIP. When people are projected to extend or reenlist, their numbers can be added to the extendee or reenlisted array. However, when extensions and reenlistments occur which are represented by retention rates (with actual levels determined by COMPLIP) they must be tracked in a separate array, called "extensions and reenlistments of retention rates." The portion of the array for 3 and 4 year initial terms of enlistment is represented in Figure 2. Unlike the numbers representing people, retention rates cannot be added to the reenlisted array.

The array in Figure 2 is large: there will be a maximum of 40 C-groups and 7 years for each t. In addition, at least 89 classes in the month-of-accession dimension will be required. As the projection proceeds past month $\pi(2)$, the first month for which COMPLIP determines FT RA accesions, the array will begin to fill with retention rates, $r_{i,j}$, for the C-groups/terms $j = 1,...n$. The non-zero-filled rows will run from $\pi(2)$ to i, the current projection month. In its current implementation, this will be a 16 by 7 by 89 array.

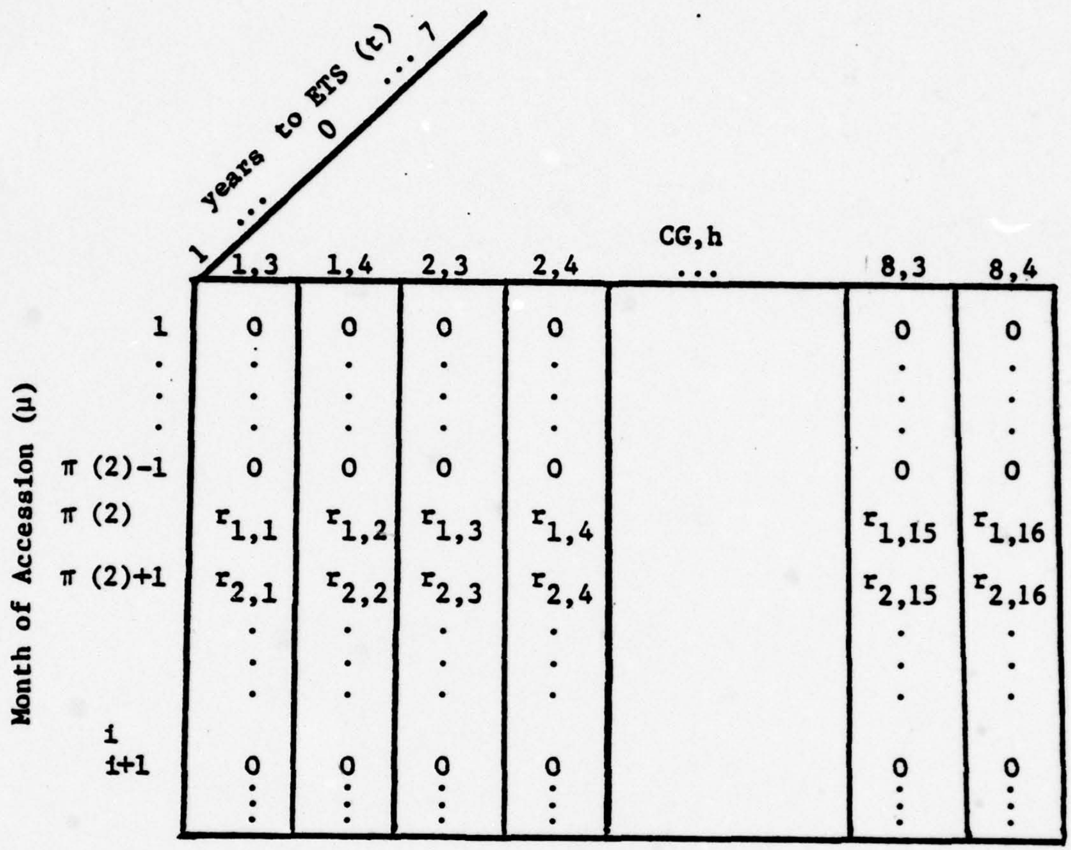


Figure 2. Extension Retention Rate Array

Modifications to Other Modules

The routines in DPM/QDPM which establish historical inventories for arrays not classified by C-group--AUS, FTI, STI, CAR, and the three area arrays--will be changed very little. They will be modified to assign people to the appropriate arrays based to the criteria discussed above.

Modifications Affecting COMPLIP

Since C-groups will include term of service, 40 sets of retention rates might have to be input to COMPLIP to compute NPS gains for each C-group. Although technically feasible, this would substantially increase COMPLIP's run time. To avoid such a large increase in the size of the linear program, the C-groups for terms 3 and 4 are summed across term in the IPM as are the C-groups for terms 2, 5, and 6. Factors are provided by QFAC to spread NPS accession reguirements by C-group back to the specific terms of enlistment in the Report Generator.

APPENDIX A

THE ELIM-IV QUALITATIVE DATA BASE

INTRODUCTION

This appendix discusses the enhancement to the data bases which will be used to provide C-group expansion and tracking, including extendee tracking. This would be added to those of ELIM-II and ELIM-III to accomplish the following tasks:

- Expand C-groups from 4 to 40, with term of service as a C-group characteristic
- Maintain data by C-group for a longer period of service
- Track enlisted personnel until separation or first reenlistment, explicitly separating first from career reenlistments (relates to "first enlistment cohort tracking" and "first-timer to reenlistee transition" in proposal)
- Simplify redefinition of C-groups and
- Enable user to calculate historical loss rates and current inventories by marital status

It is also designed to facilitate projecting C-group transitions, if that task is undertaken in the future.

A tracking file has been created with one record for each individual in the data base. The record indicates his/her entry and, where potentially different, current characteristics, and the month and type of each transaction relevant to ELIM-COMPLIP. The small numeric data base has been expanded to 40 C-groups; each month it can be updated directly or it can be re-created from the new tracking file. There is no more need for an intermediate large numeric qualitative data base. The system allows for an individual to change C-groups. Persons would be counted in C-groups until discharge.

DESCRIPTION OF THE DATA BASE

The new data base, the tracking file, exists in two forms: the so-called large (Table A.1) and small (Table A.2) tracking files. They replace

TABLE A.1

LARGE TRACKING FILE

1-9	Social Security Number	
10	NPS Indicator	
11-16	Date of Birth	
17-18	AFQT	
19	Sex	
20	Race	
21-24	Term of Service	
22-25	BPED	
26	Education (Initial)	
27	Mental Category	
28-31	Enlistment Options	
32-33	CMF	
34	Pay Grade (Initial)	
35	Eligibility to Reenlist	
36	Moral Waiver	
37	Character of Separation	
38	Prior Service Indicator	
39-42	ETS	
43-45	UIC	
46	Bonus	
47	Times Reenlisted	
48-50	Age at Entry, months	
51	Ethnic	
52	Receiving Station (no good)	
53-56	BASD	
57	Marital Status	
58	Number Dependents	
59-60	Initial Assignment	
61-63	Initial Location	
64-66	Initial Primary MOS	
67	Current Pay Grade	
68-69	Current Assignment	
70-72	Current Location	
73	Current Education	
74-76	Current Duty MOS	
77-79	Current Primary MOS	
80-81	Current Status Code	
82-83	No. of Trailers	
T*1	... T*24	
84	268	Type Transaction[1]
85-88	269-272	Date
89-91	273-275	Transaction Code[2]

---

[1]

1	Gain-Loss	5	Education
2	Pay Grade	6	Duty MOS
3	Assignment	7	Primary MOS
4	Location	8	Status Code

[2] For EMF variables (type transaction 2-8), new value of variable; for GLF, 3 character SPN or SPD if loss; EBE, etc. if extension; GRF, GHI, etc., for gains.

TABLE A.2

SMALL TRACKING FILE

1-4	Cohort
10-13	Social Security Number
14-15	AFQT
16	Race
17	Sex
18	Term Of Service
19	Civilian Education
20	Mental Category
23-25	Age at Entry, months
21-22	CMF
26-29	BPED
30-33	ETS
34-36	UIC
37	(Recv)
38-41	BASD
42-43	Current Assignment
44-46	Current Duty MOS
47-49	Current Primary MOS
50-54	Flags Initial EMF data present Trailers lost (overflow) Extension Loss Return
55-56	Number of Trailers
57-128	Trailers: 57 Type[1] … 123 Type[1] 58-59 Month Service … 124-25 Month Service 60-62 Transaction … 126-28 Transaction

[1]currently, always = 1

MT63, MT67, and M.68. For each individual entering the Army after 1 Jan 72, there is one record (on each version of the tracking file). This record consists of the person's Social Security Administration Number (or Temporary Identification Number), codes indicating characteristics at enlistment, codes for the current value of those characteristics which can change, and a series of transactions "trailers," one for each transaction relevant to ELIM-COMPLIP. Each trailer contains the month the transaction was recorded and the type of transaction. The large tracking file contains slightly more information on characteristics than the small. It allows for more transaction trailers to accumulate and keeps trailers for changes in characteristics as well as for gain and loss transactions; the small tracking file contains only gain-loss transactions. The large file records SPN, SPD, or other actual code; the small file aggregates into 16 categories. The small file revises ETS date when an extension or non-NPS gain occurs. The large file is in a sense an archive, from which status of Army personnel at any time in the past can be reconstructed. The small file represents current Army personnel status and is designed for efficient use in calculating loss rates by current or entry characteristics. Therefore, the large tracking file is ordered by SSAN, but the small tracking file is organized by cohort.

In each tracking file, following the SSAN (or TIN), there is information on the individual's race, sex, civilian education, etc., at the time of entering the Army. In addition, there are codes for current characteristics, since some characteristics which the Army is presently interested in can change.[1] The user can determine from the file each individual's term of service, time to ETS (including negative t), and extendee status--all especially important for the desired first reenlistment statistics.

INTEGRATING THE DATA BASE INTO ELIM-II AND ELIM-III

Some parts of the ELIM-III system will continue to operate very much as they have in the past. The main difference is that, as each month's transaction update is performed, there is no need to update MT67B, the large numeric qualitative data base. This data base currently contains historical statistics for 150 quality classes by month of service. To provide the desired flexibility for user redefinition of classes, this data array would have become excessively large (>10,000 classes). Therefore, it was replaced with the tracking file, the compressed version of the Master Transaction Data Base MT63. Each month, transaction trailers are added to records in the tracking file, gains are added to it, and current characteristic codes are changed if necessary to reflect transactions such as a change in marital status or number of dependents or

[1] Both entry and current values are maintained to provide the user the option to sort on status at either time.

education level. Then a 40[1]-class small numeric qualitative data base replacing MT68 is updated directly each month. When the user wants to redefine quality classes, he can go back to the tracking file and recreate the smaller data base. Procedures for updating the smaller data base are discussed in more detail in the next section, following the description of the data base.

MODIFICATIONS TO MT68

The "small" numeric quality data base consists of two parts. The main part stores data for first-timers; a second part contains data for extendees. Conceptually, the first-timers array of the small numeric quality data base resembles Figure A.1. The shaded box represents the data--extensions, IMRE losses, etc.--for cohort 1, C-group 1, in the first MServ. It currently consists of 11 items of information but might be expanded to 13 (see the next section, "Changing an Individual's C-Group"). There are similar sets of data for every cohort, C-group, and MServ if the cohort is old enough to have served that number of months.

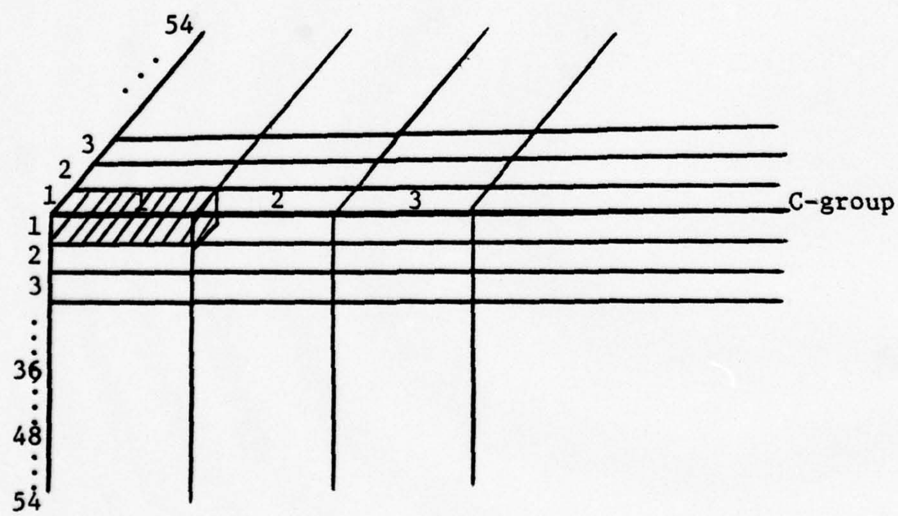


Figure A.1. Expanded MT68

[1]Maximum.

All members of a given C-group have the same term of service, h. Although time to ETS, t, is not explicit in the data, it is completely defined by h and MServ. For the older (earlier) cohorts, there are non-zero data not only in MServ cells 1 through $12 \cdot h$, but also in cells with MServ $>12 \cdot h$, for individuals with $t<0$.

The second-timer array of the small numeric data base will carry second-timers through loss from the Army or for 54 months. The data in it will be used for producing historical data, not for inventory projections.

One of the characteristic codes in the tracking file indicates whether the individual is a first-timer, an extendee, a reenlistee, or a previous loss. During each monthly update, as each transaction is added to the tracking file, this characteristic code is checked to determine the relevant array for the affected individual. If he/she is a first-timer, the appropriate C-group-MServ-cohort in the first-timer array of the small numeric data base is manipulated. In addition, if the transaction is an extension, reenlistment, or loss, the individual is added to the extendee or "careerist" array, if appropriate, and the characteristic code for the storage array is updated. The "careerist" array becomes a reenlisted array. Because an individual can reenlist before completing 36 months of service, this array is modified to cover MServ <37. In addition, the ELIM-III DPM is modified so that individuals are assigned to this array only when they reenlist. Currently an individual can be assigned to the careerist array if he/she is an extendee with MServ >36; if such a person subsequently reenlists, his/her first reenlistment appears to be a career reenlistment.

If the transaction affects an individual in the extendee array, the appropriate cell in that array is updated and the other procedures described for first-timers would be executed. If the affected individual is a reenlistee, ELIM-II is already accounting for him/her and the only action taken will depend on the type of loss and will parallel procedures in the current system.

With separate data arrays for first-timers, extended first-timers, and reenlisted personnel, it is possible to calculate reenlistment rates separately for first and for subsequent reenlistments. This is not possible in ELIM-III.

Since first timers are to be carried by cohort, no explicit monthly aging will be needed; it will be automatic. Extendees will be carried by months to ETS, which must be reduced by 1 each month (unless a new extension is executed).

Each person remains in the small numeric data base for first timers until the earliest of (a) separation from service (loss); (b) IMmediate REenlistment loss; (c) extension; or (d) completion of 54 months of service. The 54-month cutoff allows for reenlistment of all persons with 4-year first enlistments and up to a 6-month extension. Extendees are carried until they reenlist or separate. Reenlistees are carried in a careerist-type array.

The small numeric data base (C-group arrays) is saved each month (as MT68B is now) and used as input (MT68A) the next month. In fact, if the user is experimenting with various sets of C-groups, several versions of the C-group arrays can be saved until a final determination as to the desired characteristics is made.

CHANGING AN INDIVIDUAL'S C-GROUP[1]

In each monthly update of the data, individuals can be classified on the basis of the current, not entry, characteristics. Thus, an individual need not be in the same C-group every month. For example, if an individual were single from MServ = 1 until a marriage in MServ = 28, he/she would be included in the inventory of a C-group for single ... persons with each monthly update until MServ 28 and thereafter in a different C-group, one for married ... persons (with "..." signifying otherwise similar characteristics).

---

[1]This capability is not provided under the current contract; but it will be possible to include it in ELIM-IV, with a modest additional effort.

This necessitates the introduction of an additional loss and an additional gain category, thus expanding the data in MT68 from 11 to 13 words per cohort, C-group, and MServ, as noted above. That is, each cell, as represented by the shaded area in Figure A.1, would be expanded to make room for a ninth loss category, transition to a different C-group, and a new gain category, transition from a different C-group. If changes in marital status were the only transitions to be considered, C-groups for single persons could have losses due to marriage and gains due to divorce (if divorced and single status are deemed equivalent); C-groups for married persons could have transitional gains due to marriage and losses due to divorce.

Source and destination C-groups would not be linked; that is, an individual could move from a C-group to any other C-group. There would not be a set of C-groups for single individuals and a parallel set for married individuals with one-to-one links between them.

When the data base is updated, the presence of a transaction involving a C-group transition will produce a loss in the individual's beginning-of-month C-group and a gain in his/her end-of-month C-group. If a set of redefined C-groups is to be created anew, routines will have to be devised that keep track of C-group for the MServ being dealt with at that stage in the data creation. This C-group need not necessarily be the same as the entry or the current--i.e., most recent--C-group.

INVENTORY PROJECTIONS

Although the historical data base could accommodate changes in C-group, the Inventory Projection Module could not, without great effort, be altered to allow for changes in C-group in the projection out-months. Ideally, and eventually, before loss rates are applied to the beginning inventory for each out-month, transition loss and gain (marriage and divorce) rates would be used to reapportion the inventory among the C-groups.

On the other hand, if a characteristic which can change values, such as marital status, is used for C-group assignment and projections are made without C-group transitions in the out-months, then projections in the out-months are likely to become increasingly inaccurate, the farther into the future we project. This inaccuracy is expected because, as the inventory is aged, more and more persons will actually be married, but the projection system will not move individuals between single and married C-groups. Most NPS gain inputs to MServ = 1 will be single; and as they are aged, the projected inventory will have an unrealistically high percentage of single people.

The modifications to ELIM greatly expand the number of C-groups in the "small" numeric data base. The larger numeric data base has been replaced by a tracking file with a record for each individual. The record contains information for C-group assignment and transaction information. The data enable cohort tracking by C-group through first reenlistment, permit simplified C-group redefinition, and carry information on characteristics such as marital status that were not previously used for C-group classification.

APPENDIX B

CAREER FORCE OPTIMIZATION (CFO)

STATEMENT OF THE PROBLEM

The Army desires the capability to optimize the career strength over a specified year of service (YOS) distribution as defined by the Enlisted Force Master Plan. Gains to the career force come from several different sources, including:

- First timers with 36 months of completed service
- Reenlistees of several types
- Army and non-Army prior service enlistees with 36 months or more of prior service

Not all career gains will have the same number of years of service, therefore, it is desirable that COMPLIP be able to optimize the gains from each category, within user supplied constraints, such that the Career Force YOS distribution is driven towards, and ultimately reaches, the desired distribution within a user-specified length of time.

Currently, the Army uses a 5-year planning cycle for setting and working towards a desired YOS distribution. Within the plan, it attempts to optimize the force within the next 3 years. The current plan specifies an optimum force strength to be reached in FY 80. These are annual strengths which are determined as of 30 September of each fiscal year. Because the Army plan extends for a maximum of 5 years, there is no need to change ELIM-COMPLIP in terms of the number of years that can be forecast.

REQUIRED SYSTEM CAPABILITY

Currently, the ELIM-COMPLIP system handles monthly data and distributes the career force into two categories: retirement eligible (RE) and non-retirement eligible (NRE). To optimize a YOS distribution, it is necessary to compute strength by YOS (1,2,3,...,20,20+). In order to maintain this distribution over the period covered by the manpower program, factors are needed to age the appropriate fraction of the population month by month from one YOS to the next.

The analytical problem becomes one of minimizing the weighted sum of the deviations of the actual from the planned career force strength, by YOS, subject to constraints on total strength, the supply of NPS enlistees, the supply of PS enlistees by YOS, and the first-term reenlistment rates by years of service. (The current Army policy is to exercise control only at the first reenlistment. Beyond the first reenlistment, the Army generally keeps all people who desire to reenlist. However, the system will have the capability, through user input, to override later reenlistment rates so that the user may experiment with alternative reenlistment policies.)

METHODOLOGY

The development of the CFO methodology was separated into two phases. In Phase I, the YOS distribution for the career force was developed and in Phase II, the harder problem of optimizing the various gains to the career force to meet a specified career force profile has been attacked.

Phase I

The methodology for this phase involved changes to the DPM, FAC, IPM, COMPLIP equations, and the report generator to produce a YOS distribution of enlisted strength. This has been discussed in previous concept papers.

Phase II

Solution of this phase represents one of the most difficult tasks of any in the history of ELIM-COMPLIP development. The following sections discuss the alternatives considered and the alternative used for solution of the task.

The Objective Function

From the preceding discussion, it is clear that the objective function should model the Army's goal of reducing the deviation between the career force strength[1] and a target profile. Specifically, minimize

---

[1] Although the term "career force strength" will be used for clarity throughout this paper, the optimization will minimize the weighted difference between actual and target strength for YOS 1, 2, and 3 as well as the career force.

the weighted sum of the absolute values of the differences between projected year-end trained strength of the career force by years of service and target values (Career Force Profile) in the target year. It is not necessary to have additional profiles in each of the years preceding the target year since the target profile at the end of year .0X will drive all accessions that will end up in that profile. Target profiles may not be needed before the target year but may be desirable. Profiles after the target year are certainly needed.

An alternative to making the user create a target profile for each year of projection is to use the <u>same</u> profile as in the target year but relax the acceptable deviations more, the further away the projection year is from the target year. That is, if

FCZ = fraction of the target that represents acceptable deviations

RCF = factor for adjusting FCZ for other than target years, RCF > 1

NYR = target year

IYR = projection year

then, define FC(IYR) as

$$FC(IYR) = FCZ * RCF^{|IYR-NYR|}$$

Note that FC(IYR) = FCZ when IYR = NYR

$$\overline{FCTARijj} = FC(IYR) * \overline{CTARGnjj} \; ; \quad i = IYR, n = NYR$$

where $\overline{CTARGnjj}$ is the target profile in the target year n and $\overline{FCTARijj}$ represents the acceptable deviations in projection year i. Graphically, FC(IYR) is as shown in Figure B.1.

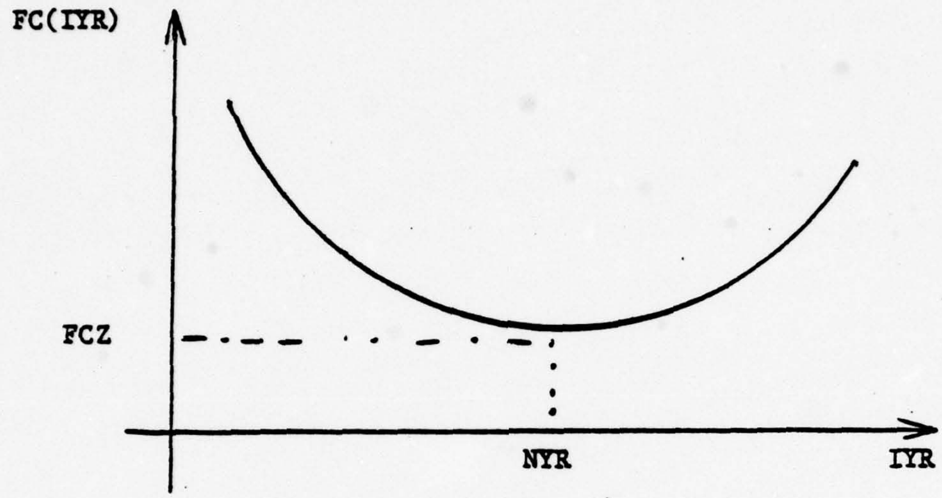


Figure B.1. Graph of FC(IYR)

The user would provide the desired force distribution, the year by which it is to be met, the weights for the difference between YOS strength and target for each YOS, and the factor which moves the force towards that target. The formulation for the objective function is presented in Annex 1 of this appendix.

The concept for the objective function follows from the problem statement; however, the concepts for varying reenlistment gains are not as straightforward.

Because most of the gains to the career force enter with 3 and 4 years of service, there is very little that can be done in the short term to adjust current deviations beyond YOS 10. However, this is not a significant problem, for two reasons: (1) approximately two-thirds of the absolute value of the deviations from the current force profile occur within YOS 1 through 10, with the greatest deviations being in YOS 4 through 7, and (2) beyond YOS 10, the retention rates are quite steady. This means that if the curve for YOS 1 through 10 can be adjusted in the short-term, the corrections to the rest of the curve should come automatically with time.

Varying First Reenlistments

To model a CFO decision process, first reenlistments must be allowed (within specified limits) to vary independently of NPS gains. However, the number of first reenlistments by a given cohort is actually the product of factors representing attrition and reenlistment rates and a variable representing the number of NPS gains for that cohort. If the reenlistment rate is replaced by a variable, then the number of reenlistments becomes the product of two variables--a nonlinear function which is nonconvex. Nonconvex functions generally preclude both linear programming and the efficient solution methods for convex nonlinear problems. Conceptually, they are in a category with the integer problems which lack a network structure; problems of this type with more than six to ten variables are often considered "large" and may require several hours or more to solve. Therefore, it is clear that it is not reasonable simply to add additional variables, representing reenlistment factors, to the present COMPLIP constraint set.

Since the idea of replacing the linear program in COMPLIP by a nonlinear program is infeasible, other alternatives must focus on the treatment of the terms representing reenlistment gains. (The term "reenlistment" in subsequent discussions will always imply "first reenlistments.") Nonlinearity can be avoided by modeling reenlistment losses as they currently are, that is, as linear functions of NPS gains.

Iterating Between Successive LPs

One approach is to develop a linear program which would only vary the reenlistment factors, with NPS gains being fixed. In operation, COMPLIP would, using the current LP, find a solution vector of NPS gains which optimized the CFO objective function with reenlistment rates fixed;

then the same objective function would be solved by another LP formulation which holds NPS gains fixed at the previous "optimal" value and solves for a factor of reenlistment factors. These factors would become constants in the "original" LP and it would be solved again for a new set of NPS gains. The iteration would continue until there was negligible improvement in the objective function.

This technique avoids dealing with nonlinear programs but the iterative process raises other problems. There is no way of knowing how many iterations would be required for convergence, or if the process would converge at all. Convergence characteristics may be a function of weighting factors in the objective function or combinations of values for the technology coefficients. Quite possibly the operator would have no method of forecasting convergence. Other difficulties include the mechanics of communication between the two LPs and the increased maintenance load which such a complex routine could place on support personnel. A solution scheme not requiring iterations would be more desirable.

Separating Reenlistment Gains and Losses

Another method to avoid nonlinearity is to disconnect the computation of reenlistment gains and losses. Reenlistment losses would remain functions of NPS gains and the appropriate factors; however, reenlistment gains would become a new COMPLIP variable which would be upper and lower bounded by a user-supplied linear function of reenlistment losses. In effect, this would permit the first termers and career force to vary independently (within reasonable limits). The basic COMPLIP structure could be retained as all variables would be linear and solution iterations would not be required. COMPLIP would not be modified other than those changes necessary to add additional variables and constraint equations and an additional objective function.

Accounting for Career Force Losses

Another problem to deal with is the computation of losses to the populations that have reenlisted. Heretofore, the post-reenlistment losses, and thus the contribution to the inventory of those who have reenlisted, were modeled in the IPM by extension/reenlistment retention rates. These rates were passed to COMPLIP and became technology coefficients representing the reenlisted population as a proportion of NPS gains. This relationship would not be maintained under the above scheme since the "optimal" reenlistment factors would not be known until completion of the COMPLIP run.

One approach would be to create an additional set of rates for the IPM to represent the losses to the reenlisted population. COMPLIP does not require career losses by type of loss, it simply needs to know how many post-reenlisted are present in the inventory. Given the number of reenlisted (actually a variable representing reenlisted), aggregate factors could be used to model the attrition of the reenlisted in COMPLIP. This would be accurate enough for COMPLIP's requirements but would not provide information on the various categories of losses for the Report Generator. In other words, once the correct numbers of reenlistments have been selected, there must be a projection of losses by type.

Losses by type could be determined by constructing this capability in the Report Generator. Actually this would require building a sort of IPM. It would seem more logical to use the loss projection capability that already exists--in the IPM. This suggests that after COMPLIP determines the optimum reenlistment rates (as well as NPS gains, etc.), we return to the IPM to allocate the career losses. Since NPS gains are known, the IPM would produce an "all people run" (that is, retention rates would not be used), losses would be correctly allocated, and the results would proceed to the Report Generator. To format these results for the Report Generator, it will be necessary for the solution to pass via the COMPLIP solution file, but the LP solution would be immediate.

This approach solves another problem, that of properly allocating reenlistments. Although feasible, it would appear unwieldy and very burdensome to the operators for COMPLIP to compute different reenlistment factors, for each projection month, for each of 40 C-groups.

Since the Army does not goal reenlistments by C-groups, and the cost in computer time for COMPLIP to model variable reenlistments for each C-group could easily add several hours to a COMPLIP run, we propose to have COMPLIP determine an aggregate reenlistment factor for all C-groups (and all terms of service). Then, after COMPLIP is solved, but before returning to the IPM for the second pass, a new processor would allocate the aggregate reenlistment rate over all C-groups in proportion to the original rates. The second pass to the IPM would then allocate correctly the reenlistments by C-groups. This scheme for assigning reenlistments to C-groups would preserve the proportion of reenlistments dictated by reenlistment losses unless the user desires otherwise.

Prior service gains will continue to be treated in the same manner as in the current system.

The CFO equations and additional comments on the methodology are contained in the annex to this appendix.

ANNEX B1

CFO COMPLIP EQUATIONS

This annex contains the basic COMPLIP equations to be implemented for career force optimization (CFO).

COMPLIP EQUATIONS FOR CAREER FORCE OPTIMIZATION

OBJ...07 - objective function

$$\min . \sum_{n=1}^{Y} \sum_{jj=1}^{NYS} [W_{acp}(jj)*ACPOSnjj+W_{acn}(jj)*ACNEGnjj$$

$$+ W_{exp}(jj)*EXPOSnjj+W_{exn}(jj)*EXNEGnjj]$$

where,

multiplication is denoted by *

n is the target year

NYS is the maximum number of YOS categories

y is the number of years in manpower program

jj is the year of service (YOS) index

ACPOSnjj is the variable for the number of acceptable positive deviations from target in year n for YOS jj

ACNEGnjj - acceptable negative deviations

EXPOSnjj - excessive positive deviations

EXNEGnjj - excessive negative deviations

W_{acp} - OBJ...07 weight for ACPOSnjj

W_{acn} - OBJ...07 weight for ACNEGnjj

W_{exp} - OBJ...07 weight for EXPOSnjj

W_{exn} - OBJ...07 weight for EXNEGnjj

E.CFDnjj - CF deviations from target

$$CF.Y.njj - ACPOSnjj-EXPOSnjj+ACNEGnjj+EXNEGnjj = \overline{CTARGnjj}$$

for n=1,2,...,Y; jj=1,2,...NYS

where,

CF.Y.njj - denotes the career force (CF) strength in target year n with YOS jj

$\overline{\text{CTARGnjj}}$ - denotes the CF strength in year n with YOS jj

Note: Currently NYS = 21 and YOS 21 includes YOS > 20

UCDACnjj - upper limit on acceptable deviations

$$\text{ACPOSnjj+ACNEGnjj} \leq \overline{\text{FCTARnjj}}$$

for n=1,2,...,Y; jj=1,2,...,NYS

where,

$$\text{FCTARnjj} = \text{FC(n)*}\overline{\text{CTARGnjj}}$$

FC(n) = user-supplied factor for computing acceptable limit of deviations from target

C.OBJ.07 - constraint on OBJ...07

$$\sum_{n=1}^{Y} \sum_{jj=1}^{NYS} [W_{acp}\text{(jj)*ACPOSnjj}+W_{acn}\text{(jj)*ACNEGnjj}$$

$$+ W_{exp}\text{(jj)*EXPOSnjj}+W_{exn}\text{(jj)+EXNEGnjj}] \leq \overline{v}_7$$

where,

$\overline{v}_7$ = slightly larger than the OBJ...07 functional

E.CFYnjj - CF strength in target year n with YOS jj

$$\text{CF.Y.njj} - \sum_{\ell\ell=3}^{6} \sum_{\substack{kk\in P_{\ell jn} \\ j=jj}} \text{IRG.}\ell\ell\text{kk*}\eta_{IRG}\text{(kk,12n)}$$

$$- \sum_{qq=1}^{NCG34} \sum_{\substack{kk \in P_{\ell jn} \\ j=jj \\ \ell=0}} FTQCqqkk*\eta_{qq}(kk,12n)$$

$$- \sum_{\substack{kk \in P_{\ell jn} \\ j=jj \\ \ell=\ell\ell}} FTRTT.kk*\eta_{tt}(kk,12n)$$

$$- \sum_{qq=1}^{NCG256} \sum_{\substack{kk \in P_{\ell jn} \\ \ell=0 \\ j=jj}} FTQ2qqkk*\eta_{qq2}(kk,12n) = \overline{CFP.Ynjj}$$

where,

CF.Y.njj = CF strength in target year n with YOS jj.

IRG.ℓℓkk = IRG in COMPLIP month kk with YOSℓℓ (e.g., ℓℓ=04 for months of service 37,38,...,48).

FTQCqqkk = FT gains for CGqq in COMPLIP month k.

$\eta_{IRG}(kk,12n)$ = retention rate of IRG in COMPLIP month of accession kk at end of COMPLIP month 12n.

$\eta_{qq}(kk,12n)$ = retention rate of FT terms 3 and 4 CG qq gains in COMPLIP month kk at end of COMPLIP month 12n.

$\eta_{qq2}(kk,12n)$ = retention rates of FT terms 2, 5, 6 CGqq gains in COMPLIP month kk at end of COMPLIP month 12n.

$\overline{CFP.Ynjj}$ = CF (people not retention rates) with YOS jj at the end of year n.

$$P_{ijn} = \left\{ kk \mid j-1 < \ell\ell + n - \frac{kk-1}{12} \leq j \right\}$$

The retention rates η_{IRG} will be computed in the IPM for YOS 3, 4, 5, 6 at the time of reenlistment. These retention rates are applicable only to first reenlistments. Immediate reenlistment losses will be computed in the IPM for all data array cells for which there are non-zero immediate reenlistment loss rates. However, immediate reenlistment gains will be added only for those corresponding to second or later reenlistments. The IRG for the first timer reenlistments will be accounted for by means of the COMPLIP variable IRG.ℓℓkk.

<u>UIRGjjii</u> - upper constraint on first IRG in COMPLIP month ii with YOSjj.

$$\text{IRG.jjii} - \text{UF} \sum_{qq=1}^{\text{NCG34}} \sum_{\substack{kk \in F_{kji} \\ jj=j=3,4 \\ i=ii \\ k=kk}} \text{FTQCqqkk} * \eta_{\text{IRL}}(qq,kk,ii)$$

$$-\text{UF} \sum_{qq=1}^{\text{NCG256}} \sum_{\substack{kk \in F_{kji} \\ jj=j=5 \\ i=ii \\ k=kk}} \text{FTQ2qqkk} * \eta_{\text{IRL2}}(qq,kk,ii)$$

$$\leq \text{UF} * \overline{\text{PIRLjjii}}$$

<u>LIGRjjii</u> - lower constraint on first IRG in COMPLIP month ii with YOSjj.

$$\text{IRG.jjii} - \text{LF} \sum_{qq=1}^{\text{NCG34}} \sum_{\substack{kk \in F_{kji} \\ jj=j=3,4 \\ i=ii \\ k=kk}} \text{FTQCqqkk} * \eta_{\text{IRL}}(qq,kk,ii)$$

$$-\text{LF} \sum_{qq=1}^{\text{NCG256}} \sum_{\substack{kk \in F_{kji} \\ jj=j=5 \\ i=ii \\ k=kk}} \text{FTQ2qqkk} * \eta_{\text{IRL2}}(qq,kk,ii)$$

$$\geq \text{LF} * \overline{\text{PIRLjjii}}$$

where,

UF and LF are user supplied factors defining the first IRG constraint band width (e.g., UF=1.05, LF=.95).

$\eta_{IRL}(qq,kk,ii)$ = fraction of FT CGqq for terms 3 and 4 with month of accession kk who are IR losses in COMPLIP month ii.

$\eta_{IRL2}(qq,kk,ii)$ = fraction of FT CGqq (for terms 2, 5, 6) with month of accession kk who are IR losses in COMPLIP month ii.

$\overline{PIRLiijj}$ represents IRL from the "people" inventory that are in the first reenlistment window in COMPLIP month ii with YOS jj.

<u>E.ENLSii</u> - enlisted strength computation

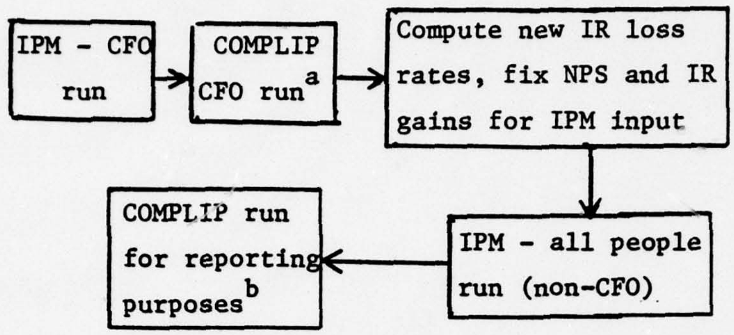
$$ENLS..ii - \Big|_{kk=ii-1} ENLS..kk - NPSG..ii+AUSL..ii + FTL...ii$$

$$- \sum_{\text{all } jj} IRG.jjii+ \sum_{\substack{kk=\pi(2)-\alpha+1 \\ \text{all } jj}}^{ii-1} [\eta_{IRG}(kk,ii-1)-\eta_{IRG}(kk,ii)]\ IRG.jjkk$$

$$= \overline{NGLii}$$

for ii = A_E+1,...,M+5.

Following is a simple schematic for Career Force Optimization:



[a]In this pass COMPLIP can have three objective functions optimized in sequence, e.g.,

OBJ...02 - operating strength optimization
OBJ...07 - career force optimization
OBJ...03 - REP optimization

[b]In this pass only OBJ...03 is used.

APPENDIX C

USER MODIFICATIONS OF LOSS FACTOR PROJECTIONS BY OPERATIONS ON LOSS DATA AGGREGATED BY ACCESSION COHORT

BACKGROUND

In earlier versions of ELIM, inventories, gains, and losses are tracked by month of service for first-timers through their first 21 months. Loss rates are projected by extension of time series of data by month of service. No attempt is made to modify loss rates in a given month of service due to changes in loss rates for earlier months of service for that same accession cohort. This factor projection and modification capability is in conflict with manpower planning procedures and OSD guidance where acceptable attrition is now being defined on a cumulative basis by accession cohort.

DATA AVAILABILITY

The revised ELIM-IV qualitative data base carries transactions out to at least 48 months of service and will be constructed for ease and efficiency of retrieval. All the historical data required for construction of cohort retention curves over months of service is in the new data base.

GENERAL METHODOLOGY

In ELIM-III, loss rates are projected and displayed as a monthly time series for a particular month of service (and particular C-group) as in Figure C.1.

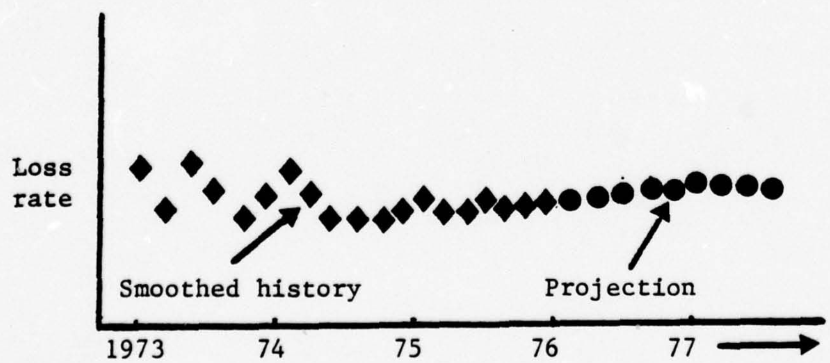


Figure C.1. Historical and Projected Loss Rates for Month of Service "X," C-Group "Y," and Loss Type "Z"

Given a complete set (all months of service, all loss types for a given C-group) of data like Figure C.1, it is possible to create the cohort-oriented plots of Figure C.2. While the transition to a plot of a single month cohort is obvious, it is also probably impracticable. However, given a set of monthly factors (great accuracy not required) for accession levels, these data can be easily aggregated up to annual or semiannual cohorts.

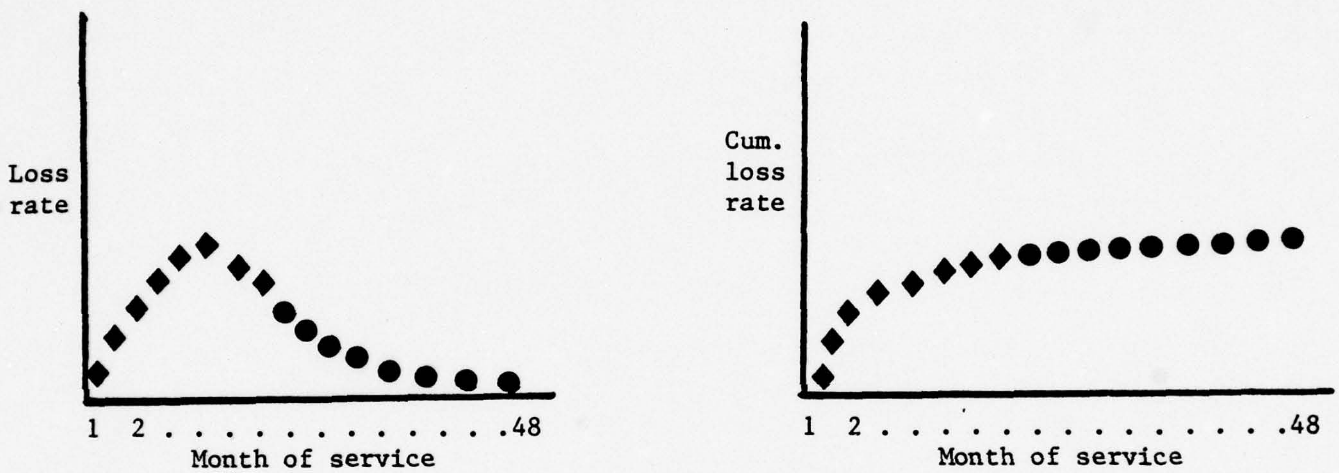


Figure C.2. Historical and Projected Loss Rates for Cohort "X" and C-Group "Y"

These plots in Figure C.2 are in terms that are more useful for projection purposes when losses are constrained by accession cohort. The Cohort Targeting System provides simple user modification tools for these types of data. The user can (iteratively) modify either or both forms. For the conditional loss rates on the left, the user specifies the month of service range and the percent reduction in the loss rate (Figure C.3). For the cumulative loss rates on the right, the user specifies the cumulative rate at an end point, probably 36 or 48 months of service, and the month of service range over which required changes in loss rates are to be made to achieve the designated target loss rate (Figure C.4).

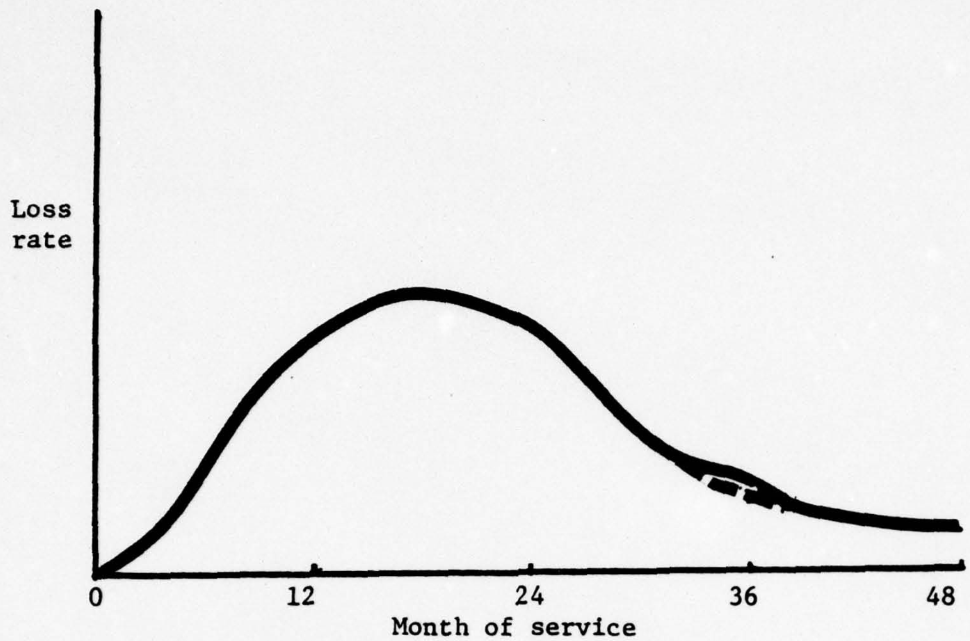


Figure C.3. Projected Loss Rates with Losses in Months of Service 30-36 Reduced by Some Specified Percentage

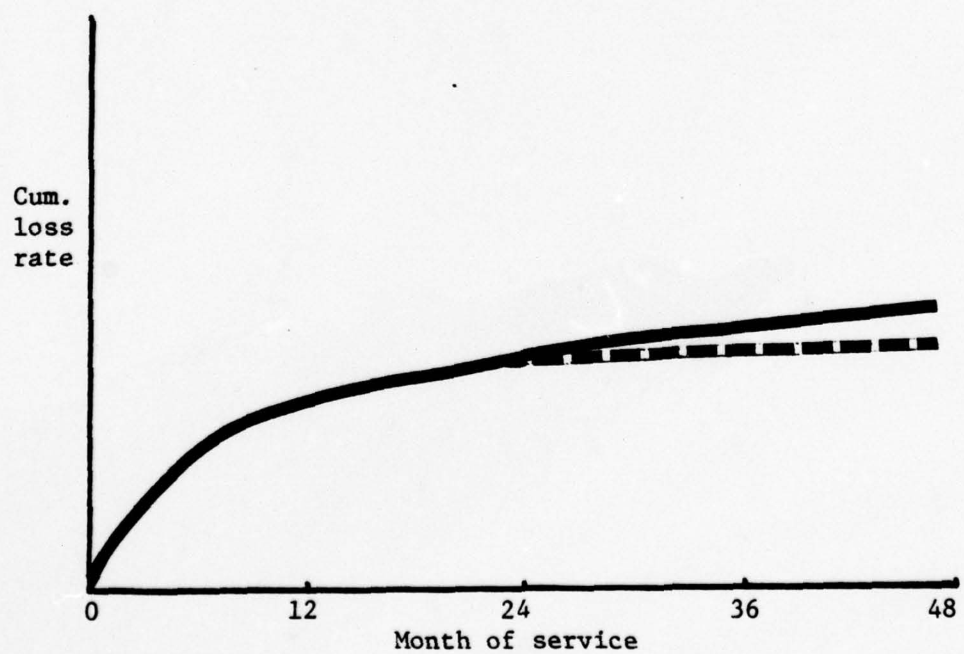


Figure C.4. Projected Cumulative Loss Rate with Losses in Months of Service 24-48 Reduced to Achieve a Percentage Reduction in the Cumulative Loss Rate at 48 Months of Service

Once the user has arrived at versions of the loss projections by cohort as in Figure C.2, these modifications are disaggregated back to the form required by the QIPM, i.e., as in Figure C.1.

OPERATIONAL CONSIDERATIONS

For several reasons, this capability exists independent of the current (or planned) production flow. That is, its use is not required for successful operation of ELIM-IV. This program operates interactively, so that the user can conveniently select a cohort, see the loss patterns that QFAC is passing to QINV, and iteratively modify the cohort loss rates to accomplish a convergence to the pattern which satisfies all the users' constraints. It is also possible to get hard copy outputs so that the user does not have to make all modifications in a single pass.

Nevertheless, under either mode of operation the principal consideration still is that this capability has been constructed as an optional enhancement to the normal procedure, that the input from this program uses the basic QFAC projections, and that the output is in the form of revised QFAC projections (at least to the extent that the user wishes to modify the QFAC outputs as a result of his cohort-based loss analysis).

AN INITIAL COMMAND SET

To implement this system, a program has been written to accept user commands and act on them. The specifications for the primary commands are provided in the following pages.

```
++++++++++++++++++++++++++++++++++++++++++++
+++++ COHORT TARGETING SYSTEM COMMANDS +++++
++++++++++++++++++++++++++++++++++++++++++++

   COHORT -- SELECTS USER-SPECIFIED SECTION OF DATA

   MODIFY -- AFFORDS USER MEANS BY WHICH TO MANIPULATE DATA

  DISPLAY -- PRODUCES GRAPH(S) OF USER-SPECIFIED SECTIONS OF DATA

     KEEP -- ALLOWS USER TO STORE DATA FOR FUTURE RETRIEVAL

     LOAD -- AFFORDS USER ACCESS TO PREVIOUSLY STORED DATA

  REPLACE -- REPLACES ORIGINAL DATA WITH MODIFIED DATA

 HARDCOPY -- ALLOWS CREATION OF PRINT-FILE CONTAINING GRAPHS AND ALL
             ACCEPTED COMMANDS, AND USER-GENERATED NOTES

   PARAMS -- PROVIDES INFORMATION PERTINENT TO CURRENT COHORT

    SCALE -- ALLOWS USER TO SPECIFY FREQUENCY OR CUMULATIVE SCALE TO
             BE USED FOR GRAPHING PURPOSES

TERMINATE -- CAUSES CTS SESSION TO END

(HELP)  ? -- PROVIDES USER WITH INFORMATION REGARDING COMMAND
             FUNCTION, SYNTAX, AND OPERANDS

    QUERY -- LISTS KEEP-SETS IN USE AND THEIR LABELS, STATUS OF
             HARDCOPY, AND CURRENT VALUES OF CUMULATIVE & FREQUENCY
             SCALES

     NOTE -- ALLOWS USER TO RECORD NOTES BOTH ON TERMINAL AND TO
             PRINT-FILE

     FREE -- RELEASES USER-SPECIFIED KEEP-SETS FOR RE-USE
```

```
FUNCTION:  THE COHORT COMMAND ALLOWS THE USER TO SELECT A
           SPECIFIC SECTION OF DATA FOR MODIFICATION OR
           EXAMINATION.

  SYNTAX:  COHORT 'YYMM'('NMON') TERM('T1',...'T5') CG('G1',...)

   ALIAS:  C

OPERANDS:  'YYMM' -- THE YEAR AND MONTH OF DATA AT WHICH TO BEGIN.

           'NMON' -- AN INTEGER VALUE SIGNIFYING THE NUMBER
                     OF MONTHS OF DATA DESIRED.

           TERM -- ALIAS: T
               SPECIFIES THAT ONE OR MORE TERMS ARE DESIRED.
               'T1' -- INTEGER SIGNIFYING A TERM.

           CG -- SPECIFIES THAT ONE OR MORE CHARACTERISTIC
                 GROUPS IS DESIRED.
                 'G1' -- INTEGER SIGNIFYING A
                         CHARACTERISTIC GROUP.

    NOTE:  ALL OPERANDS ARE REQUIRED.
```

```
FUNCTION:  THE MODIFY COMMAND ENABLES THE USER TO MAKE CHANGES
           TO DATA AS DESIRED.

  SYNTAX:  MODIFY CUMULATIVE MONTHS('M1'/'M1','M2')
         - TARGET('MTGT','TVAL')

      OR:  MODIFY FREQUENCY MONTHS('M1'/'M1','M2')
         - PERCENT/CONSTANT('PCVAL')
         - LOSSTYPES(ALL/'TYPE','TYPE',...'TYPE')

OPERANDS:  CUMULATIVE -- ALIAS: C
               SPECIFIES THAT CUMULATIVE DATA IS TO BE MODIFIED.

           FREQUENCY -- ALIAS: F
               SPECIFIES THAT FREQUENCY DATA IS TO BE MODIFIED.

           MONTHS -- ALIAS: M
               SPECIFIES THE RANGE OF MONTHS (INCLUSIVE)
               OF DATA TO BE MODIFIED.
               'M1' -- INTEGER SIGNIFYING THE FIRST MONTH.
               'M2' -- INTEGER SIGNIFYING THE LAST MONTH.

           TARGET -- ALIAS: T
               INDICATES THAT TARGET-MONTH AND TARGET-VALUE
               FOLLOW THIS OPERAND.
               'MTGT' -- INTEGER SIGNIFYING THE TARGET-MONTH.
               'TVAL' -- REAL-NUMBER SIGNIFYING THE TARGET-VALUE.
                         DECIMAL POINT IS REQUIRED.

           PERCENT/
           CONSTANT -- ALIAS: P/C
               SPECIFIES THAT MODIFICATION OF DATA WILL OCCUR
               BY MEANS OF EITHER A PERCENTAGE OR CONSTANT
               VALUE, DEPENDING UPON WHICH FORM OF THE OPERAND
               IS ENTERED, PERCENT OR CONSTANT (P OR C).
               'PCVAL' -- REAL-NUMBER SIGNIFYING THE VALUE OF THE
                          PERCENT OR CONSTANT EMPLOYED TO MODIFY
                          THE DATA.  DECIMAL POINT IS REQUIRED.

           LOSSTYPES -- ALIAS: L
               SPECIFIES THAT LOSS-TYPES TO BE MODIFIED FOLLOW
               THIS OPERAND.
               ALL -- SPECIFIES THAT ALL LOSS-TYPES ARE TO BE
                      MODIFIED.
               'TYPE' -- AN ALPHABETIC ENTRY SIGNIFYING A LOSS-TYPE
                         TO BE MODIFIED, TAKEN FROM AMONG THE
                         FOLLOWING:
                                     ETS
                                     IMRE
                                     DFMC______________
                                     TRDP   EITHER FORM
                                     TDP ___ACCEPTABLE.
                                     MCDT
                                     UNFT
                                     PHYD
                                     UNKN
                                     OTHR
                                     ADMN
```

EXTN

LOSSTYPE ENTRY MAY INCLUDE ANY NUMBER OF TYPES DESIRED, AND MAY BE ENTERED IN ANY ORDER.

NOTE: IF THE CUMULATIVE OPERAND IS SPECIFIED, MONTHS AND TARGET MUST ALSO BE SPECIFIED. THE CUMULATIVE OPERAND MUST IMMEDIATELY FOLLOW THE COMMAND-NAME, MODIFY, OR THE COMMAND-ALIAS, M. THE MONTHS AND TARGE OPERANDS MAY BE SPECIFIED IN ANY ORDER.

IF THE CUMULATIVE OPERAND IS SPECIFIED, THE PERCENT/CONSTANT AND LOSSTYPES OPERANDS MAY NOT BE SPECIFIED.

IF THE FREQUENCY OPERAND IS SPECIFIED, THE MONTHS, PERCENT/CONSTANT, AND LOSSTYPES OPERANDS MUST ALSO BE SPECIFIED. THE FREQUENCY OPERAND MUST IMMEDIATELY FOLLOW THE COMMAND-NAME, MODIFY, OR THE COMMAND-ALIAS, M. THE MONTHS, PERCENT/CONSTANT, AND LOSSTYPES OPERANDS MAY BE SPECIFIED IN ANY ORDER.

IF THE FREQUENCY OPERAND IS SPECIFIED, THE TARGET OPERAND MAY NOT BE SPECIFIED.

```
  FUNCTION:  THE DISPLAY COMMAND INITIATES THE PRODUCTION OF GRAPHS
             DEPICTING EITHER FREQUENCY DISTRIBUTIONS OR CUMULATIVE
             DISTRIBUTIONS, OVER 48-MONTH PERIODS, OF USER-SPECIFIED
             SECTIONS OF DATA.

    SYNTAX:  DISPLAY CUMULATIVE

        OR:  DISPLAY FREQUENCY(ALL/'TYPE','TYPE',...'TYPE')

  OPERANDS:  CUMULATIVE -- ALIAS: C
                 SPECIFIES THAT A GRAPH OF A CUMULATIVE DISTRIBUTION
                 IS DESIRED.

             FREQUENCY -- ALIAS: F
                 SPECIFIES THAT A GRAPH OF A FREQUENCY DISTRIBUTION
                 IS DESIRED.
                 ALL -- SPECIFIES THAT A GRAPH IS DESIRED OF EACH
                        LOSS-TYPE.
                 'TYPE' -- AN ALPHABETIC ENTRY SIGNIFYING A LOSS-TYPE
                           FOR WHICH A GRAPH IS DESIRED, TAKEN FROM
                           AMONG THE FOLLOWING:
                                                  ETS
                                                  IMRE
                                                  DFMC______________
                                                  TRDP   EITHER FORM
                                                  TDP ___ACCEPTABLE.
                                                  MCDT
                                                  UNFT
                                                  PHYD
                                                  UNKN
                                                  OTHR
                                                  ADMN
                                                  EXTN

                           LOSS-TYPE ENTRY MAY INCLUDE ANY NUMBER
                           OF TYPES DESIRED, AND MAY BE ENTERED IN
                           ANY ORDER.

      NOTE:  IF CUMULATIVE IS SPECIFIED, FREQUENCY AND LOSS-TYPES
             MAY NOT BE SPECIFIED.

             IF FREQUENCY IS SPECIFIED, LOSS-TYPES MUST BE SPECIFIED
             AND CUMULATIVE MAY NOT BE SPECIFIED.
```

FUNCTION: THE KEEP COMMAND ALLOWS THE USER TO SAVE MODIFIED VERSIONS OF ORIGINAL DATA. UP TO 19 SETS OF MODIFIED DATA MAY BE KEPT. CTS AUTOMATICALLY EXECUTES A KEEP COMMAND WHENEVER THE USER ISSUES A COHORT COMMAND.

SYNTAX: KEEP 'NUM' ('LABEL')

ALIAS: K

OPERANDS: 'NUM' — INTEGER SIGNIFYING THE KEEP-SET WHERE THE MODIFIED VERSION OF THE DATA IS TO BE STORED. MUST BE IN THE RANGE: 21 > NUM > 0.

'LABEL' — UP TO 64 CHARACTERS MAY BE ENTERED FOR THE PURPOSE OF USER-DIFFERENTIATION BETWEEN VERSIONS OF MODIFIED DATA FOR A SINGLE COHORT.

NOTE: NEITHER 'NUM' NOR 'LABEL' ARE REQUIRED, THOUGH EITHER OR BOTH MAY BE SPECIFIED. IF 'NUM' IS NOT SPECIFIED, CTS WILL ASSIGN THE LARGEST PREVIOUSLY UNUSED KEEP-SET NUMBER. IF ALL KEEP-SETS HAVE BEEN USED, CTS WILL GIVE THE USER AN OPPORTUNITY TO RE-USE A KEEP-SET OF THE USER'S CHOICE. IF 'LABEL' IS NOT SPECIFIED, CTS WILL ASSIGN A LABEL OF "NONE". THE LABEL FOR THE KEEP-SET WHICH IS USED BY CTS UPON USER-ISSUANCE OF A COHORT COMMAND WILL CONTAIN THE COHORT COMMAND-STRING ITSELF, AS ENTERED BY THE USER.

```
FUNCTION:  THE LOAD COMMAND RESTORES USER-ACCESS TO DATA WHICH
           HAS BEEN PREVIOUSLY SAVED BY MEANS OF A KEEP COMMAND.

  SYNTAX:  LOAD 'NUM'

   ALIAS:  L

OPERANDS:  'NUM' -- INTEGER SIGNIFYING THE KEEP-SET FROM WHICH
                    DATA IS TO BE RETRIEVED.  MUST BE IN THE
                    RANGE:  21 > NUM > 0.

    NOTE:  'NUM' IS REQUIRED.  IF NOT SPECIFIED, OR IF OUT OF
           RANGE, CTS WILL REJECT THE COMMAND.
```

```
FUNCTION:  THE REPLACE COMMAND ALLOWS THE USER TO SUBSTITUTE
           MODIFIED DATA IN PLACE OF THE ORIGINAL DATA.

  SYNTAX:  REPLACE 'YYMM'

   ALIAS:  R

OPERANDS:  'YYMM' -- THE MONTH AND YEAR AT WHICH TO BEGIN
                     REPLACEMENT OF ORIGINAL DATA BY MODIFIED
                     DATA.

    NOTE:  'YYMM' IS A REQUIRED ENTRY.
```

```
 FUNCTION:  THE HARDCOPY COMMAND ALLOWS THE USER TO CONTROL THE
            WRITING OF GRAPHS AND ACCEPTED COMMANDS TO THE PRINT
            FILE.

   SYNTAX:  HARDCOPY ON/OFF/ONLY

    ALIAS:  H

 OPERANDS:  ON -- CAUSES GRAPHS TO BE WRITTEN BOTH ON THE TERMINAL
                  AND TO THE PRINT-FILE.
            OFF -- CAUSES GRAPHS TO BE WRITTEN ON THE TERMINAL ONLY,
                   AND NOT TO THE PRINT-FILE.
            ONLY -- CAUSES GRAPHS TO BE WRITTEN TO THE PRINT-FILE
                    ONLY, AND NOT ON THE TERMINAL.

     NOTE:  THE SETTINGS ARE MUTUALLY EXCLUSIVE--ONLY ONE FORM IS
            ACCEPTABLE PER ISSUANCE OF THE HARDCOPY COMMAND.
```

```
FUNCTION:  THE PARAMS COMMAND PROVIDES THE USER WITH THE FOLLOWING
           INFORMATION:

           1.  CURRENT YEAR AND MONTH

           2.  YEAR & MONTH OF FIRST COHORT IN HISTORY FILE

           3.  NUMBER OF MONTHS OF HISTORY

           4.  NUMBER OF MONTHS OF PROJECTION

           5.  NUMBER OF CHARACTERISTIC-GROUPS, 2-YR TERM
                                                3-YR TERM
                                                4-YR TERM
                                                5-YR TERM
                                                6-YR TERM

  SYNTAX:  PARAMS

   ALIAS:  P

OPERANDS:  NONE
```

```
  FUNCTION:  THE SCALE COMMAND PROVIDES THE USER THE OPTION OF
             SPECIFYING THE FREQUENCY AND CUMULATIVE SCALES TO
             BE USED IN THE CONSTRUCTION OF GRAPHS OF DATA, OR
             TO SET SCALES OFF (VALUE=0.0).

    SYNTAX:  SCALE FREQUENCY/CUMULATIVE OFF/'YMAX'

     ALIAS:  S

  OPERANDS:  FREQUENCY -- ALIAS: F
                  SPECIFIES THAT THE SETTING SPECIFIED IS TO BE
                  APPLIED TO GRAPHS GENERATED USING FREQUENCY DATA.

             CUMULATIVE -- ALIAS: C
                  SPECIFIES THAT THE SETTING SPECIFIED IS TO BE
                  APPLIED TO GRAPHS GENERATED USING CUMULATIVE DATA.

             OFF -- SPECIFIES THAT A SCALE-FACTOR IS TO BE RESET
                      TO THE VALUE OF ZERO.

             'YMAX' -- A REAL-NUMBER SIGNIFYING THE SCALE-FACTOR
                           DESIRED.  MUST BE IN THE RANGE: 1.0 > YMAX > 0.
                           DECIMAL POINT IS REQUIRED.

      NOTE:  EITHER FREQUENCY OR CUMULATIVE MUST BE SPECIFIED.
             BOTH MAY NOT BE SPECIFIED.

             EITHER OFF OR 'YMAX' MUST BE SPECIFIED.
             BOTH MAY NOT BE SPECIFIED.
```

MODIFY COMMAND IMPLEMENTATION

Frequency Array Modification

given: range of months of service, $m_1 \rightarrow m_2$
loss type
percent or constant change

(1) Make specified changes.
(2) Check to see if next command is another frequency array modification. If Yes, go to 1; if No, go to 3.
(3) Recompute frequencies for "ALL" losses.
(4) Recompute cumulative loss array.

Cumulative Vector Modification

given: range of months of service, $m_1 \rightarrow m_2$
Target loss rate, T, and associated Month of Service, K

(1) Make changes as follows:
- Find $S_i = 1 - CUM_i$ for all $i \geq m_1$
- Find $\left(S'_{m_2}\right) = S_{m_2}\left(S_K/(1-T)\right)$
- Find $$F = \frac{1 - \left(\dfrac{S''_{m_2}}{S_{m_1}}\right)^{1/(m_2-m_1+1)}}{1 - \left(\dfrac{S_{m_2}}{S_{m_1}}\right)^{1/(m_2-m_1+1)}}$$
- Find $S_i = S_{i-1}(1-F) + S_i F$: for $m_1 \leq i \leq m_2 - 1$
- Find $C_i = 1 - S_i$ for all $i \geq m_1$

(2) Recompute frequencies for "ALL LOSSES"
(3) Recompute frequencies for individual loss types, excluding ETS losses

APPENDIX D

STRUCTURE SPACES AS VARIABLES

STATEMENT OF THE PROBLEM

The Army desires the capability to treat structure spaces as linear programming variables. This capability would allow the analyst to control the input of structure spaces, or targets, within the matrix generator, and subsequently in the LP.

REQUIRED SYSTEM CAPABILITY

When optimizing on OBJ1 or OBJ2, COMPLIP minimizes the weighted sum of the absolute values of the deviations from the target structure spaces. OBJ2 has two parts, acceptable and excessive deviations. The excessive deviations are weighted more heavily. When making the target structure spaces a variable in each projection month, they will adjust themselves--within the upper and lower bounds--to make the deviations even closer to zero.

When the range of the lower and upper bounds is sufficiently great in each month, all the deviations would likely be computed as zero by adjusting the variable target, hence computing an "idealized" target--idealized in the sense of minimum deviations. Whether this "idealized" target is useful in the context of the manpower program is another question.

It is also conceivable that with the lower and upper bound range sufficiently large, COMPLIP would adjust the variable target to either extreme in order to minimize deviations rather than adjust the accessions. This should certainly be checked by making experimental runs. The user must determine acceptable limits on the variable target in order to produce a reasonable manpower program and adjusted structure spaces.

METHODOLOGY

The following section presents the source code modifications to the matrix generator which were delivered to USAMSSA to be implemented and tested by them.

In Subroutine EDIT the user table TARG must be re-defined to contain two additional columns with column names FAC1 and FAC2. Column FAC1 has for each month the fraction of the target in column RHS (the only column now in TARG) that becomes the lower bound on the new variable TAR...ii. Similarly, the column FAC2 contains the values for the upper bound. Typical values in FAC1 and FAC2 would be .95 and 1.05 respectively. That is, values in FAC1 are less than or equal to 1.0, and in FAC2 they are greater than or equal to 1.0.

Let t represent the index in array TNAME for table TARG in EDIT. Then the following change is in order:

```
DATA NCOL (t) / 3 /
```

In MATGEN, the sections of the code corresponding to E.AVD1ii and E.AVD2ii need to be changed. The code for U.DAC.ii need not be changed.

```
C     E.AVDIII ***
C IF  (FOUND. . .
      IF (.NOT. (FOUND (4HOBJ1,. . .
      CALL FETCH (4HTARG, ARRAYU, . . .
      JOFSET = 3+JORG(1, LIM1, LIM2)
      KOFSET = 3+JORG(0, LIM1, LIM2)
      J2OFST = 3+JORG(2, LIM1, LIM2)
      J3OFST = 3+JORG(3, LIM1, LIM2)
      CALL ENDTAR(ARRAYU(KOFSET+1), ARRAYU(JOFSET+1,M,LIM2)
      CALL ENDTAR(ARRAYU(KOFSET+1),ARRAYU(J2OFST+1,M,LIM2)
      CALL ENDTAR(ARRAYU(KOFSET+1),ARRAYU(J3OFST+1,M,LIM2)
      CALL EAVD1(ARRAYU(JOFSET+1),ARRAYU(J2OFST+1),
     * ARRAYU(J3OFST+1),A1,LIM2)
      . . .
```

for *** E,AVD2II *** the coding changes are similar:

```
    .
    .
    .
    J2OFST = . . .
    J3OFST = . . .
    CALL ENDTAR            . . .      JOFSET+1, . . .
         "                 . . .      J2OFST+1, . . .
         "                 . . .      J3OFST+1, . . .
    CALL EAVD2(ARRAYU(JOFSET+1),ARRAYU(J2OFST+L),ARRAYU(J3OFST+1),
         * A1,LIM2)
```

The changes to subroutines EAVD1 and EAVD2 are exactly the same. Each has three line changes, the calling sequence, a type statement and the CALL RHS2(. . .). That is:

```
    SUBROUTINE EAVD {1/2} (TAR,FAC1,FAC2,A1,M5)
    . . .
    REAL TAR(M5),FAC1(M5),FAC2(M5)
    . . .
```

<u>replace CALL RHS2 (. . .) with</u>

```
    CALL COL22(4HTAR.,2H..,I,4HE.AV,2HD{1/2},I, * -TAR(I))
    CALL BND2(2HLO,4HTAR.,2H..,I,FAC1(I))
    CALL BND2(2HUP,4HTAR.,2H..,I,FAC2(I))
```

where {1/2} means: use 1 when EAVD1 is changed and use 2 when EAVD2 is changed.

APPENDIX E

MATRIX GENERATOR ENHANCEMENTS

STATEMENT OF THE PROBLEM

The Army desires the capability within the matrix generator to express linkages between the various characteristic groups as ratios or percentage of each other (e.g., Mental-Category-IV High-School-Diploma-Graduate equals 10 percent of all High-School-Diploma-Graduate Accessions during a particular projection period). In addition to this enhancement, the capability to equate one of the characteristic groups to WAC females and to accommodate the special constraints of the budget year by developing different capabilities for the budget year and the remaining years will be new capabilities within the matrix generator.

REQUIRED SYSTEM CAPABILITY

The existing ELIM-COMPLIP system has been modified to provide the user with the capability within the matrix generator to specify the accession level of first-term enlistments of one characteristic group as a fraction of its accession level of another characteristic group and the capability to equate one of the characteristic groups to user-supplied WAC females.

The design and source code modifications necessary to incorporate these new features into COMPLIP have been made and delivered to USAMSSA to be implemented and tested by them. Once COMPLIP has been positively tested, the modified matrix generator code must be incorporated with the modifications made by GE TEMPO for the training model.

METHODOLOGY

The following represent the source code modifications which allow linkages between the various characteristic grops, one characteristic group to be equated to WAC females, and the special constraints for the budget year as new capabilities of the matrix generator.

<u>In MGEDIT</u>

UPDATE <u>TOBJ</u> to accommodate type 4 tables with two digit numerals up to the number 24 (e.g., CG 01, 02, __, 16; CG 21, 22, 23, 24). <u>Skip 17, 18, 19, 20</u>.

New User Tables

Proportionality Constraints Relating Group 1 to Group 2

TABLE PROF

	FAC1	FAC2	...	FAC6
MONFY				
.				
.				
.				

Proportionality Constraint set 1

GRP1 = FAC1 * GRP2

TABLE PRO1

	GRP1	GRP2
01	1.0	0.0
02	1.0	0.0
.	.	.
.	.	.
.	.	.
16	1.0	0.0
21	0.0	1.0
22	0.0	0.0
23	1.0	0.0
24	0.0	0.0

Constraint set 2

GRP1 = FAC2 * GRP2

TABLE PRO2

	GRP1	GRP2
01	1.0	0.0
02	1.0	0.0
.	.	.
.	.	.
.	.	.
16	1.0	0.0
21	0.0	1.0
22	0.0	0.0
23	1.0	0.0
24	0.0	0.0

TABLE PRO6

e.g. PROℓ(GRP1)ii=PROF(FACℓ)*PROℓ(GRP2)

<u>COMPLIP Equation</u>

EYROPℓii

$$-\sum_{jj\in P_{2\ell}}^{NCG34} FTQCjjii - \sum_{jj\in P_{2\ell}}^{NCG256} FTQ2jjii + f_{ii1} \sum_{\substack{jj=1 \\ jj\in P_{2\ell}}}^{NCG34} FTQCjjii + f_{ii1} \sum_{\substack{jj=1 \\ jj\in P_{2\ell}}}^{NCG256} FTQ2jjii = 0$$

Where $P_{1\ell}$, $P_{2\ell}$ are defined as follows:

$P_{1\ell}$= set of Characteristic Groups that are to be proportional to a second set $P_{2\ell}$; the proportionality constant being $f_{ii\ell}$ e.g. P_{11} refers to column GRP1 in Table PRO1, P_{21} refers to column GRP2 in Table PRO2, and f_{ii1} refers to FACI for month ii in Table PROF.

Code Modifications

MATGEN: At front in MATGEN

```
INTEGER PROP(6)
DATA PROP/4HPRO1, 4HPRO2, 4HPRO3, 4HPRO4, 4HPRO5, 4HPRO6/
```

Insert the following just before "CALL CURTIS(...)":

```
C     *** EPROP ***    TWO SETS OF CG PROPORTIONAL  TO EACH OTHER
      DO  2190 L=1,6
      IF (.NOT. (FOUND(PROP(L), DIRU, LDIRU))) GOTO 2190
      FETCH (UHPROF, ARRAYU, LNU, BUFU, LBUFU, DIRU, LDIRU,
             LUNU, MSGFIL)
      JOFSET = 3+JORG (L, LIM1, LIM2)
C     move proportionality constant into WORK1
      CALL MOVER (WORK4(LIMI), ARRAYU (JOFSET + LIM1), 1 - LIM1 + LIM2)
      LIM1F = LIM1
      LIM2F = LIM2
      DO  2090 II = LIM1F, LIM2F
      IROW = L*100 + II
      CALL ROW3 (1HE, 4HEPRO, 1HP, IROW)
      FETCH (PROP(L), ARRAYU, LNU, BUFU, LBUFU, DIRU, LDFRU, LUNU, MSGFIL)
      JOFSET = 3 + JORG(L, LIM1, LIM2)
      KOFSET = 3 + JORG(2, LIM1, LIM2)
C     CG 1-16 TERMS 3,4
      DO  2060 ICG34 = 1, NCG34
C     ARAYU IS A REAL ARRAY EQUIV TO THE INTEGER
C     ARRAY ARRAYU
      IF (ARAYU (JOFSET + ICG34) . LE.0.0)GOTO 2050
      CALL COL23 (4HFTQC, FTQC2 (ICG34), II, 4HEPRO, 1HP, IROW, -1.0)

2050  IF (ARAYU (KOFSET + JCG34) .LE.0.0)GOTO 2060
      CALL COL23 (4HFTQC, FTQC2 (ICG34), II, 4HEPRO, 1HP, IROW,
                - WORK4 (II))

2060  CONTINUE
C
C     CG 21, 22, 23, 24,  TERMS 2,5,6
```

APPENDIX F

AN OVERVIEW OF CURRENTLY AVAILABLE STATISTICAL FORECASTING MODELS

Prepared for the General Research Corporation by
Warren Rogers Associates, Inc., 12 Sunnyside Place,
Newport, Rhode Island

APPENDIX F
AN OVERVIEW OF CURRENTLY AVAILABLE FORECASTING MODELS

Summary

This paper provides an overview of statistical forecasting models. The objective is to provide an understanding of the strengths and limitations of what is currently available in this field rather than an analytical treatment of any specific model. While the various general classes of models are mentioned, the major emphasis is on time series extrapolative models as these would seem most appropriate for the forecasting problem being addressed by General Research. Annex F1 contains bibliography of the relevant literature.

Quantitative forecasting models have received considerable attention in recent years; analytic endeavors can be stratified into three primary groups: survey and market research methods, time-series (extrapolative) methods, and causality methods. (A fourth group might include methods oriented towards the systematic development of a consensus of expert opinion; the Delphi method is most familiar in this regard.) The three data-based groups have been developed for use in rather distinct problem settings; although there is occasional debate regarding model choice across the groupings, most frequently models within a single group are suggested as appropriate in any given problem.

Survey and market research methods of forecasting are most generally applied to problems in which doubt exists as to the existence of a credible data base. Typical applications of such methods would encompass such decisions as new product introduction or the outcome of an election. Implicit in the choice of these methods is the assumption that the problem under study is sufficiently unique to warrant independent analysis. Survey and market research methods closely parallel standard methods of statistics, classical or Bayesian. They involve the formulation of the problem (e.g., the unknowns to be estimated or hypotheses to be tested), experimental design (e.g., what types and quantity of data to collect), actual data collection, and analysis.

Forecasting is a direct outcome of such endeavors, in the sense that the entire effort is directed towards the relevant "unknowns" in the decision environment. The primary drawbacks of methods within this broad category are expense and their "one-time" value. As a result, typical applications are generally restricted to non-recurring problems of significant consequence.

Causal models are applied to problems in which a dependent variable (or a set of dependent variables simultaneously) is believed related to a set of independent variables, with random errors intruding on the underlying structural relationship. The initial impetus for the development of models within this category was provided by various economic forecasting problems. Most familiar of the causal models is regression analysis, in which a dependent variable, y, is believed linearly related to a set of independent variables, $X_1, \ldots, X_n$, by the relation

$$y = B_0 + B_1X_1 + \ldots + B_nX_n + U,$$

where the B's are unknown coefficients and n is an unobserved error term. This single-equation model is readily extended to simultaneous-equation versons within which a set of dependent variables, $y_1, \ldots, y_m$ are jointed determined within a set of m equations, each potentially involving both dependent (y) variables and independent (x) variables as well as an error term. Forecasting with causal models involves two stages of analysis: the interpretation of the unknown parameters of the model from data on the y's and x's, and the preparation of forecasts from the estimated model. Subsidiary analytic efforts are often required in the latter stage, for example, when the independent variables themselves are not known with certainty. Forecast accuracy with causal models are governed by three interrelated factors. The quality of the estimates of the unknown model parameters, the relative magnitudes of the structural and random components of the model, and the degree of certainty regarding the inputs to the model when it is used for forecasting.

The econometric literature has supplied a wide variety of statistical proceedings for most causal models of both the single and simultaneous equation types. Given that a causal model can be postulated (i.e., the relevant independent variables identified) and the existence of a credible data base, standard procedures exist for both model estimation and forecasting. These procedures generally allow inferences to be drawn about both the adequacy of the model in explaining historical data and the accuracy of forecasts made using the model. When causality exists and is identified, models drawn from this general class are clearly appropriate. Efforts of this type are generally more durable than those previously discussed; while considerable one-time expense may be incurred in formulating and estimating causal models, they are easily updated for repeated application.

The third category of analytic models, time-series or extrapolative models, applies to problems in which future values of a variable are to be predicted using prior observations on the same variable as a basis. Extrapolative models, as well as being important in their own right, are often substituted for causal models. Among the earliest applications of these models was to the problem of sales forecasting. This application suggests an important caveat in the use of extrapolative models. Sales, in most cases, can be said to be "caused" by such factors as pricing policy, advertising, etc. Using previous sales figures to forecast future sales may be reasonable if pricing and advertising policy have remained unchanged; on the other hand, past sales figures may be a poor guide to the future when the causing variables are changing.

The recent spurt of development efforts regarding extrapolative models has stemmed from an attempt to retain the parsimony inherent in this class of models while avoiding the need to implicitly assume an unchanging causal environment. The former feature is clearly one worth retaining: extrapolative models generally require the minimum attention for repeated application, drawing only upon the variable of interest for updating and forecasting. Only the initial effort of choosing among extrapolative models is noteworthy in terms of expense or difficulty. The latter factor -- the true causal environment -- is the key drawback to models within this category. Extrapolative models cannot be expected to perform well when conditions change dramatically. A corollary to this drawback is that those models can be expected to perform best in short-term forecasting, since extrapolative forecasts of the distant future implicitly assume a stable environment.

The remainder of this paper discusses various extrapolative models. Section 2 outlines the simplest such models and expands upon the discussion of this section. Section 3 extends these models to consider various additional regularities in data which such models can accomodate. Section 4 describes more advanced methods developed in recent years which incorporate most of the earlier efforts; these models require initial analytic efforts, but retain the quality of parsimony once developed. Finally, Section 5 presents a summary and suggestions for consideration in the present application.

2. Simple Extrapolative Models

In this section, the general estimation problem to be considered is that of determining the future values of some variable x from a time series of observations, x_1, x_2, ..., x_T, on the past values of this same variable. The observations

$x_1, \ldots, x_T$ are ordered chronologically, beginning with the value observed in the earliest period, x_1, and progressing to the most recently observed value, x_T.

The most frequently employed extrapolative estimation techniques fall generally into the category of weighted averages of the previous history of observations. A weighted average forecast of the variable x, denoted $\tilde{x}$, based upon the series of observations $x_1, x_2, \ldots, x_T$ may be written as

$$\tilde{x} = w_1 x_1 + w_2 x_2 + \ldots + w_T x_T ,$$

where $w_1, w_2, \ldots, w_T$ are weights applied to the previous observations. Typically these weights are constrained to sum to one:

$$w_1 + w_2 + \ldots + w_T = 1.$$

Simple and moving average estimators both fall within this general class. A simple average estimator is one in which the weight assigned to each past observation is equal to 1/T. Thus the simple average estimator is

$$\bar{x} = (1/T)x_1 + (1/T)x_2 + \ldots + (1/T)x_T$$

$$= \frac{x_1 + x_2 + \ldots + x_T}{T}$$

$$= \frac{1}{T}\sum_{i=1}^{T} x_i \quad .$$

A moving average estimator of length N assigns weights of (1/N) to the most recent N observations and weights of zero to the earliest (T-N) observations. Thus a moving average estimator of length N is

$$\tilde{x} = (0)x_1 + \ldots + (0)x_{T-N} + (1/N)x_{T-N+1} + \ldots + (1/N)x_T$$

$$= \frac{x_{T-N+1} + \ldots + x_T}{N}$$

$$= \frac{1}{N} \sum_{i=T-N+1}^{T} x_i \quad .$$

For example, a moving average estimator of length N=3 would be defined as .

$$\tilde{x} = \frac{x_{T-2} + x_{T-1} + x_T}{3}$$

To examine the properties of these two estimators, consider the following simple process by which the actual values of the observations might be generated:

$$x_i = a + u_i \ ,$$

where a is an unknown constant and u_i is a random disturbance with a mean of zero (i.e., $E(u_i) = 0$) and a variance of σ^2 (i.e., $\text{Var}\ (u_i) = \sigma^2$) that affects each period's observation independently. The observations, x_i, will thus tend to cluster around the value a, with the spread determined by the variance of the random disturbance. For this simple process, it can be shown that both the simple and moving average estimators defined above have expected values equal to the constant a (i.e., $E(\tilde{x}) = a$); thus, on average, both will tend towards the deterministic component of the process generating the data. The variances of the two estimators, however, differ. For the simple average estimator, $\text{Var}(\tilde{x}) = \sigma^2/T$, while for the moving average estimator of length N, $\text{Var}(\tilde{x}) = \sigma^2/N$. Since the length of the full series exceeds the length of any moving average estimator (T>N), the variance of the simple average estimator will be smaller. From these calculations, the general inference might be drawn that a simple average estimator dominates a moving average estimator when the underlying process generating the observations is stable from period to period (other than an unpredictable component such as the disturbance u_i).

This conclusion, however, is altered when the underlying process is less stable. For example, if for some reason the constant value, a, shifts to a different value, b, after some time t=M in the process (i.e.,

$$x_i = \begin{cases} a + u_i & \text{for} \quad t \leq M \\ b + u_i & \text{for} \quad t > M, \end{cases}$$

where the disturbances u_i are as before), then it can be shown that the expected values of the simple and moving average estimators are no longer the same. The expected value of the simple average estimator can be determined to be

$$E(\tilde{x}) = \frac{Ma + (T-M)b}{T} .$$

Thus, future estimates using the simple average estimator will be forever biased away from b as a result of the inclusion of the data from the first M periods. A moving average estimator of length N, on the other hand, will have as its expected value $E(\tilde{x}) = b$ after period M+N (i.e., after the transitional period). Therefore, the general inference may be drawn that the moving average estimator adjusts better than the simple average estimator to changes in the underlying process.

The above simple examples suggest two generally desirable properties of extrapolative estimators: the ability to adjust to structural changes that occur in the underlying process and stability under the condition that observations may merely reflect random disturbances rather than changes in the actual process. As suggested above, the moving average estimator is preferable with regard to the first property and the simple average is preferable with regard to the second.

However, since both the presence of random disturbances and the possibility of structural change are typically present in most environments to which such estimation models might be applied, neither of these estimators is entirely satisfactory.

A third weighted average technique, known as exponential smoothing, combines the concepts underlying the two averaging techniques discussed above in that it both places weight upon all the past observations in making forecasts (as does the simple average) and weights more heavily data from the most recent periods (as do moving average models). The exponential smoothing technique places weight $w_i = \lambda(1-\lambda)^{T-i}$, where λ is a number between zero and one, on each observation x_i and thus can be written as

$$\tilde{x} = \lambda(1-\lambda)^{T-1}x_1 + \lambda(1-\lambda)^{T-2}x_2 + \ldots + \lambda(1-\lambda)x_{T-1} + \lambda x_T \quad .$$

The weights applied to the observations thus decline geometrically from the most recent observation to the earliest observation.

As the length of the series of past observations grows, the sum of the weights approaches one. In fact, since for any reasonable large value of T the weights applied to early observations are small (e.g., for T=20 and λ=0.5, the weight w_1 applied to the first observation is 0.00000095), the time series will be considered to be arbitrarily long. The exponential smoothing equation can then be re-written, using summation notation, as

$$\tilde{x} = \sum_{i=0}^{\infty} \lambda(1-\lambda)^i x_{T-i} \quad .$$

This equation can be re-written in several useful ways. First, by separating the term applying to the most recent observation (i.e., λx_T), the remaining terms can be rewritten as

$$\sum_{i=1}^{\infty} \lambda(1-\lambda)^i x_{T-i}$$

$$= (1-\lambda) \sum_{i=1}^{\infty} \lambda(1-\lambda)^{i-1} x_{T-i}$$

$$= (1-\lambda) \sum_{i=0}^{\infty} \lambda(1-\lambda)^i x_{T-1-i}$$

$$= (1-\lambda)\tilde{x}_{-1},$$

where $\tilde{x}_{-1}$ is the exponential smoothing forecast after T-1 observations (but before the T^{th} observation). The exponential smoothing equation can then be rewritten as

$$\tilde{x} = \lambda x_T + (1-\lambda)\tilde{x}_{-1} \quad ,$$

i.e., the predicted value $\tilde{x}$ is a weighted average of the previous prediction $\tilde{x}_{-1}$ and the most recent observation x_T. The choice of a smoothing constant (λ) close to zero places greater weight on the previous forecast than on the most recent observation; conversely, a value of λ close to one places relatively greater weight on the most recent data.

A simple rearrangement of the above equation suggests a third interpretation of the technique. Rewriting the equation as

$$\tilde{x} = \lambda x_T + (1-\lambda)\tilde{x}_{-1}$$

$$= \lambda x_T + \tilde{x}_{-1} - \lambda\tilde{x}_{-1}$$

$$= \tilde{x}_{-1} + \lambda(x_T - \tilde{x}_1)$$

suggests that the new forecast is obtained by revising the previous forecast $\tilde{x}_{-1}$ by some fraction λ of the error $(x_T - \tilde{x}_{-1})$ between the actual observation and the previous forecast.

The choice of a value for the smoothing constant, λ, like the choice between a simple and moving average model, depends in part upon the trade-off between stability in the presence of random disturbances and the speed of adjustment to changes in the underlying process. A low value of λ implies a preference for the former characteristic in the estimator; a higher value of λ implies a preference for the latter. In practice, one might choose a value by dividing the time series of observations into two parts, constructing exponential smoothing estimators for various values of λ on the first part of the series, and choosing the value that forecasts best (e.g., in terms of average absolute error, average squared error, (or Theil's U-statistic), on the second part of the series.

The extreme simplicity of these models allows very efficient computer implementation. Once a particular form (e.g., length of moving average or value of λ) is selected, minimal storage is required. A moving average requires storing only the most recent N observations; exponential smoothing only the most recent forecast and the current observation.

3. Inclusion of Regularities in Extrapolative Models

Many data series of interest can be decomposed into identifiable elements: a permanent component, a trend component, and a periodic (seasonal or cyclical) component. Decomposition methods allow the extraction of these components separately from the time series for use in preparing a series of forecasts. Such methods can be applied by a minor extension of the exponential smoothing models described in Section 2. The extension essentially requires only the use of three parallel smoothing models -- one for estimating each component of the model. Letting

X_t = actual observation in period t

$\overline{X}_t$ = estimated permanent component for period t

S_t = estimated periodic index for period t

A_t = estimated trend for period t

L = number of periods in a cycle, and

λ_i = smoothing coefficients,

the decomposition model can be developed as follows. First, the permanent component $\overline{X}_t$ is estimated:

$$\overline{X}_t = \lambda_1\left(\frac{X_t}{S_{t-L}}\right) + (1-\lambda_1)(\overline{X}_{t-1} + A_{t-1}).$$

The estimate is a weighted average of the observed (deseasonalized) value and the previously predicted permanent value. Next, periodic indices are updated:

$$S_t = \lambda_2\left(\frac{X_t}{\overline{X}_t}\right) + (1-\lambda_2)S_{t-L} ,$$

again as a weighted average of observation and previous prediction. Finally, trend estimates are prepared:

$$A_t = \lambda_3(\overline{X}_t - \overline{X}_{t-1}) + (1-\lambda_3)A_{t-1}$$

as a weighted average of the observed and previously predicted trends. Using the decomposed indices, forecasts for future periods can be prepared. A forecast, $X_{t,N}$, made in period t for period t+N (i.e., N periods away) can be recomposed using the equation

$$X_{t,N} = (\bar{X}_t + NA_t)S_{t-L+N} \quad .$$

Several observations can be made regarding this extension to the basic extrapolative models. First, the smoothing constants (λ's) retain their interpretation: larger values place high weight on recent observations and allow rapid response to changes in data, while smaller values provide slower response and thus greater stability. Choices of values for the λ's thus again reflect the underlying trade-off described in Section 2. Secondly, little storage is required for use of this model despite its extended coverage. Only the current observation, the current estimates of the permanent component and trend, and one cycle's work of periodic indices are required for use of the model. Rapid updating and forecasting are thus facilitated.

4. Recent Advances in Extrapolative Modelling

The classical methods described above have recently been extended by the developments of Box and Jenkins and their associates. Their contribution consisted basically of the development of efficient estimation techniques for a general class of time-series models now familiarly known as ARIMA (auto-regressive integrated moving average) models. The general application of these models is seen from a consideration of the nature of time-series data. As suggested in Section 2, an observation X can be considered to have both a premanent (stationary) component, a, and a random (stochastic) component, u:

$$X = a + u \ .$$

The ARIMA models are developed from a consideration of how a series of observations $X_1, \ldots, X_T$ can be generated from such a process. First, like the simple exponential smoothing model of Section 2, each observation could be related to the previous observarions by a linear model of the form

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + u,$$

or, in finite form,

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + u \ .$$

The above form is referred to as auto-regressive in that the equation resembles the regression models described in Section 1, with the dependent variable X_t expressed as a function of its own past values $X_{t-1}, \ldots, X_{t-p}$.

A second possible way in which the series could be generated resembles the moving average process described earlier. The actual random disturbance affecting the t^{th} observation could be written as a moving average of the form

$$U_t = V_t + \alpha_1 V_{t-1} + \alpha_2 V_{t-2} + \cdots + \alpha_q V_{t-q},$$

where the V's represent independent disturbances. The U's, however, are correlated over time and resemble a moving average.

Combining the above two models yields a process of the form

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \cdots + \alpha_{t-p} X_{t-p}$$

$$+ V_t + \beta_1 V_{t-1} + \cdots + \beta_q V_{t-q} ,$$

which is referred to as a mixed auto-regressive moving average process of orders p and q (or ARIMA (p,q)).

The final extension employed in their model can be related to the discussion of trends in Section 3. Since most time-series are not stationary, and violate the simple form $x = a + u$ upon which the above discussion was based, one can employ differencing operations to allow the series to move freely. If, for example, a simple trend exists in the data, first differences

$$Y_t = X_t - X_{t-1}$$

will be stationary, and the ARIMA general model can be applied to them. The general ARIMA model of orders p, d, and q is thus merely the ARIMA model of orders p and q applied to the d^{th} differences of the original series. (Note: the other regularity discussed in Section 3, seasonality, may be considered a special case of the correlative model described above.)

The ARIMA (p,d,q) model thus represents a most general candidate for extrapolative forecasting, one which encompasses a broad spectrum of underlying processes and allows great flexibility in terms of fitting time-series. Two statistical problems are required to be confronted in the use of such a model: the selection of the orders (p,d,q) of the process, and the subsequent estimation of the model parameters. Standard statistical programs have been developed for both phases of the analysis, and can be routinely employed to yield forecasting equations. These programs additionally yield statistics regarding goodness of fit and forecast accuracy. The choice between these methods and those described earlier is basically one of cost versus effectiveness: correctly applied, these models will certainly outperform the simpler ones; initial estimation (and to a lesser degree, regular use) will be more time consuming.

5. Conclusions

As can be seen from the foregoing, the thrust of developmental efforts in time series forecasting has been to improve reliability by incorporating such factors as underlying causal variables, seasonality, trend, non-stationarity and serial correlation. The improvements to be achieved by such modifications will naturally depend on the particular series. Their contribution in general becomes critical when long range forecasts are required.

The cost of achieving such improvements lies in the added complexity of the calculations and additional data storage. It should be noted however, that most of this added cost is the one time cost of the analysis necessary to derive the appropriate model. The increases in computer storage and running time are usually small. For example, a simple one stage exponential smoothing model could be substantially improved by a one time analysis designed to provide an initial multi-stage estimate and appropriate weighting constants. As is shown in Section 2, this improvement can be incorporated in subsequent model updates without additional calculations or storage requirements. Likewise, incorporation of seasonality and trend requires a similar one time analysis but adds only two additional variables to the data base.

Where only a few series are of interest it is generally desirable to proceed to the more sophisticated Box-Jenkins forms since the increment in regular running time and storage is small while the improvement in reliability can be great. In the case at hand, however, where approximately 20,000 series will be estimated regularly, even a small increment of either running time or storage assumes unusual significance. It would therefore be prudent to demonstrate that significant improvements are achievable before proceeding to more sophisticated modeling.

I would recommend that such one-time analyses proceed as follows. For a small subset of the series, determine whether the current model could be improved without further cost in running time or storage. Second, analyze these series for seasonality and trend and determine the additional costs and benefits of incorporating them. Third, determine the appropriate Box-Jenkins model, estimate parameters for these series, and again calculate the costs and benefits of using them.

Finally, I would recommend that consideration be given to conducting a cross-series analysis to determine how these series relate to one another. It would be very surprising if many of these series were not functionally related. A very recent development in time series analysis which is not addressed in this paper permits estimation of a series as a function of other series. The objective of such analysis would be to determine if more reliable forecasts could be achieved by the use of sophisticated models on a few series which could then be extrapolated to the rest. My intuition is that this would prove to be the case and that it would provide the best long term solution.

ANNEX F1

BIBLIOGRAPHY

1. Aigner, D. J., "A Compendium on Estimation of the Autoregressive-Moving Average Model from Time Series Data," International Economic Review, October, 1971.

2. Ayres, Robert U., 1969. Technological Forecasting and Long Range Planning, McGraw-Hill, New York.

3. Box, G. E. P., "Use and abuse of regression," Technometrics, 8, 625, 1966.

4. Box, G. E. P., and N. R. Draper, "The Bayesian estimation of common parameters from several responses," Biometrika, 52, 355, 1965.

5. Box, G. E. P., and G. M. Jenkins, "Some statistical aspects of adaptive optimization and control," Jour. Royal Stat. Soc., B24, 297, 1962.

6. Box, G. E. P., and G. M. Jenkins, "Further contributions to adaptive quality control: simultaneous estimation of dynamics: non-zero costs," Bull. Intl. Stat. Inst., 34th Session, 943, Ottowa, Canada, 1963.

7. Box, G.E.P., and G. M. Jenkins, "Mathematical models for adaptive control and optimization," A.I.Ch.E.-I. Chem. E. Symp. Series, 4, 61, 1965.

8. Box, G. E. P., and G. M. Jenkins, "Some recent advances in forecasting and control, I," Applied Stat., 17, 91, 1968.

9. Box, G. E. P., and G. M. Jenkins, "Discrete models for forecasting and control," Encyclopedia of Linguistics, Information and Control, 162, Pergamon Press, 1969.

10. Box, G. E. P. and G. M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1970.

11. Box, G. E. P., G. M. Jenkins and D. W. Bacon, "Models for forecasting seasonal and non-seasonal time series," Advanced Seminar on Spectral Analysis of Time Series, ed. B. Harris, 271, John Wiley, New York, 1967.

12. Box, G. E. P. and D. A. Pierce, "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," Journal of the American Statistical Assn., 1970.

13. Brown, R. G., Statistical Forecasting for Inventory Control, McGraw-Hill, New York, 1959.

14. Brown, R. G.: "Smoothing, Forecasting, and Prediction of Discrete Time Series," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

15. Brown, R. G. and R. F. Meyer, "The fundamental theorem of exponential smoothing," Operations Res., 9, 673, 1961.

16. Brown, T. M., "Standard Errors of Forecast of a Complete Econometric Model," Econometrica 22, 1954.

17. Butler, William F., and Robert A. Kavesh, 1966. How Business Economists Forecast, Prentice-Hall, Englewood Cliffs, New Jersey.

18. Chambers, J. C., S. K. Mullich, and D. D. Smith: How to Choose the Right Forecasting Technique, Harvard Business Review, 49:45-74 (July-August 1971).

19. Christ, C. F., Econometric Models and Methods, John Wiley and Sons, Inc., New York, 1966.

20. Cox, D. R., "Prediction by exponentially weighted moving averages and related methods." Jour. Royal Stat. Soc., B23, 414, 1961.

21. Dhrymes, P. J., Econometrics: Statistical Foundations and Applications, Harper and Row, New York, 1970.

22. Goldberger, A. S., Econometric Theory, John Wiley and Sons, Inc., New York, 1964.

23. Goldberger, A. S., A. L. Nagar, and H. S. Odeh, "The Covariance Matrices of Reduced Form Coefficients and of Forecasts for a Structural Econometric Model," Econometrica 29, 1961.

24. Grenander, U. and M. Rosenblatt, Statistical Analysis of Stationary Time Series, John Wiley, New York, 1957.

25. Hannan, E. J., Time Series Analysis, Methuen, London, 1960.

26. Helmer, Olaf, 1966. The Use of the Delphi Technique--Problems of Educational Innovations, The RAND Corporation, Santa Monica, California, December, pp. 2-3.

27. Holt, C. C., "Forecasting trends and seasonals by exponentially weighted moving averages," O.N.R. Memorandum, No. 52, Carnegie Institute of Technology, 1957.

28. Holt, C. C., F. Modigliani, J. F. Muth and H. A. Simon, Planning Production, Inventories and Work Force, Prentice-Hall, New Jersey, 1963.

29. Hood, W. C. and T. C. Koopmans, Studies in Econometric Method, Cowles Commission Monograph 14, 1953.

30. Jenkins, J. G. and D. G. Watts, Spectral Analysis and Its Applications, Holden-Day, San Francisco, 1968.

31. Klein, L. R., "An Essay on the Theory of Economic Prediction," Yrjo Johnssonin Lectures, Helsinki, 1968.

32. Koopmans, T. C., Statistical Inference in Dynamic Economic Models, Cowles Commission Monograph No. 10, 1950.

33. Leuthold, R. M., A. J. A. MacCormick, A. Schmitz and D. G. Watts, "Forecasting Daily Hog Prices and Quantities: A Study of Alternative Forecasting Techniques," Journal of the American Statistical Association, 1970.

34. McLaughlin, R. L., and J. J. Boyle, 1968. Short Term Forecasting, American Marketing Association Booklet.

35. Muth, J. F. "Optimal properties of exponentially weighted forecasts of time series with permanent and transitory components," Jour. Amer. Stat. Assoc., 55, 299, 1960.

36. Nelson, Charles R., Applied Time Series Analysis, Holden-Day, Inc., 1973.

37. Nerlove, M., "Review of Box and Jenkins," Journal of Business, 1971.

38. Pierce, D. A., "Distribution of Residual Autocorrelations in the Regression Model with Autoregressive-Moving Average Errors," Journal of the Royal Statistical Society (B), 1971.

39. Steckler, H. O.: Forecasting with Econometric Models: An Evaluation, Econometrica, 36:437-463 (1968).

40. Steckler, H. O., and F. W. Burch, 1968. "Selected Economic Data: Accuracy vs. Reporting Speed," Journal of American Statistical Association, 436-444 (June).

41. Theil, H.: "Applied Economic Forecasting," North-Holland Publishing Company, Amsterdam, 1966.

42. Theil, H., Economic Forecasts and Policy, North Holland Publishing Co., Amsterdam, 1970.

43. Theil, H., Principles of Econometrics, John Wiley and Sons, Inc., New York, 1971.

44. Wheelwright, Steven C. and Spyros Makridakis, "Forecasting with Adaptive Filtering", Revue Francaise d'Automatique, d'Informatique et de Recherche Operationelle, Winter, 1973 (Paris).

45. Wilson, G. J., "Factorization of the generating function of a pure moving average process," SIAM Jour. Num. Analysis, 6, 1, 1969.

46. Winters, P., "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science, 324-342 (April 1960).

47. Winters, P. R., "Forecasting trends and seasonals by exponentially weighted moving averages," Management Science, 6, 324, 1960.

APPENDIX G
SMOOTHING CONSTANT CONTROL
BY MONTH OF HISTORY

INTRODUCTION

In previous versions of QFAC, the parameter "MONTH" in the SC CON user control file was used to specify an accession cohort. When the user wanted to adjust smoothing constants for a particular calendar month(s) of history, he had to determine the affected month of service for each available accession cohort and adjust the smoothing constant separately for each accession cohort.

This requirement is eliminated in ELIM-IV. The user can now specify an adjustment to the smoothing constant by calendar month(s) which will automatically be applied over all relevant accession cohorts.

CARD INPUT SPECIFICATIONS

To use this expanded capability, a new set of card images should be inserted between the "*SC CON" and "9999" cards. This (optional) deck will start with:

*HIST MON

and will end with:

9998

In between will be a set of cards in the usual form to specify the following:

- "LTYPE", followed by a card with the loss type or "ALL"
- "MONLO", followed by a card with the starting calendar month for the modification, in YYMM format
- "MONHI", followed by a card with the ending calendar month for the modification, in YYMM format

- "MONI", followed by a card, in YYMM format, which will set MONLO and MONHI to the same calendar month.

- "SCVAL", followed by a card, in F10.8 format, that defines a multiplier for the smoothing constant. This multiplier is initialized to 1.0 by the program

- "DO", which calls the modifications defined by the above

There is no limit on the number of cards in this deck. The user should also note that there is no resetting of the controls after a "DO", so each of LTYPE, MONLO, MONHI and SCVAL will remain unchanged until overridden.

APPENDIX H
REVISED EQUATIONS IN THE MATRIX GENERATOR FOR THE EXPANDED C-GROUPS

INTRODUCTION

This appendix displays the equations in the Matrix Generator which have been modified for the expanded C-group capability. The development involved the treatment of terms of enlistment 3 and 4 and terms of enlistment 2, 5, and 6 as two distinct categories. The equations were modified to accommodate the number of C-groups designated for combined terms 3 and 4 (number of C-groups denoted NCG34) and the number of C-groups designated for combined terms 2, 5, and 6 (number of C-groups denoted NCG256). New equations were developed to reflect upper and lower limits pertaining to the combined terms of service (for example to constrain first term enlistments, terms of service 3 and 4 years, to be no less than a specified fraction of total enlistment); new variables and user tables were also developed to reflect C-group definition and treatment of the two term of enlistment categories.

The Matrix Generator equations in COMPLIP-G2 were "hard-wired" to "track" first-term enlistees for 21 months. The Matrix Generator equations have been modified to allow for a user-specified number of "tracking months," the maximum number of months being 54.

All revised Matrix Generator equations have been entered into the source code.

C-GROUP MODIFICATIONS AND CHANGE IN TRACKING MONTHS FROM 21 TO tt, WHERE tt IS A USER-SPECIFIED INPUT

COMPLIP-G3 VARIABLES

COLUMN NAME	INDEX RANGE	DEFINITION
FTQCjjii	ii = min $(A_E+1, A_{TE}-3),\ldots,$ M+5, jj=1,...,NCG34; optional for selected jj for ii $=\xi_K,\ldots,A_E$	FT enlistees, terms of service 3 and 4 years of characteristic group jj in month ii
FTQ2jjii	ii = min $(A_E+1, A_{TE}-3),\ldots,$ M+5, jj=1,...,NCG256; optional for selected jj for ii $=\xi_K,\ldots,A_E$	FT enlistees, terms of service 2, 5, and 6 years of C-group jj in month ii
FTRtt.ii	ii = $\pi(2)-Q+1,\ldots,M+5-tt$ FT RA remain in the Army at the end of the tt^{th} month of service following enlistment in month ii. This variable is aggregated for all terms of service (2,3,4,5,6) treated as 1 C-group	

AD-A070 451 GENERAL RESEARCH CORP MCLEAN VA OPERATIONS ANALYSIS GROUP F/G 5/9
DESIGN AND DEVELOPMENT OF ELIM-IV/COMPLIP-G3.(U)
DEC 78 P B MCWHITE, K D MIDLAM, A K BOCAST MDA903-77-C-0366

UNCLASSIFIED CR-239 NL

2 OF 2
AD
A070451

END
DATE
FILMED
8-79
DDC

COMPLIP-G3 OBJECTIVE FUNCTIONS

ROW NAME	OBJECTIVE FUNCTION	PURPOSE
OBJ...04	Minimize $\sum_{jj=1}^{NCG34} \sum_{\substack{P=A_E+1 \\ ii=P}}^{M} W_{QC}(jj,P)$ FTQCjjii + $\sum_{jj=1}^{NCG256} \sum_{\substack{P=A_E+1 \\ ii=P}}^{M} W_{Q2}(jj,P)$ FTQ2jjii	Minimizes the weighted sum of first-term enlistments for terms of service 3 and 4 years for all C-groups plus the weighted sum of first-term enlistments for terms of service 2, 5, and 6 years for all C-groups. Values of $W_{QC}(jj,P)$ and $W_{Q2}(jj,P)$ can be either positive (causing the corresponding enlistments to be minimized) or negative (causing the corresponding enlistments to be maximized) or zero (removing the corresponding enlistments from the influence of the objective function).

Constraints Required

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - REQUIRED	PURPOSE
E.ESEPii	$\text{ESEP..ii-FTL...ii-AUSL..ii} + \sum_{\substack{u=\pi(1)\\ kk=u-\alpha+1\\ \gamma=ii-kk+1}}^{ii+\alpha-1} \Delta_s(c,u,\gamma)\text{DRAFT.kk} + \Delta_s\text{FTL...ii}$	Computes enlisted separations.
	$+\ 2 \sum_{jj=1}^{NCG34} \sum_{\substack{kk=\pi(2)-\alpha+1\\ u=kk-\alpha+1\\ ii\geq\pi(2)-\alpha+1+MNSRIR-1}}^{ii-MNSRIR+1} \text{RIRQ}(ii+\alpha-1,jj,u)\text{FTQCjjkk}$	Note: Equation is written for N>1. To run N=1, dummy data will be entered (small number) for at least 1 other C-group.
	$+\ 2 \sum_{jj=1}^{NCG256} \sum_{\substack{kk=\pi(2)-\alpha+1\\ u=kk-\alpha+1\\ ii\geq\pi(2)-\alpha+1+MNSRIR-1}}^{ii-MNSRIR+1} \text{RIR2}(ii+\alpha-1,jj,u)\text{FTQ2jjkk}$	
	$+2 \sum_{\substack{kk=\pi(2)-\alpha+1\\ u=kk+\alpha-1}}^{ii-tt} \text{RIRF}(ii+\alpha-1,u)\text{FTRH.kk}$	

ROW NAME	CONSTRAINT-REQUIRED	PURPOSE

$$+\ 2 \sum_{\substack{kk=\pi(1)-\alpha+1 \\ u=kk+\alpha-1}}^{ii-MNSRIR+1} RIRD(ii+\alpha-1,u)\omega DRAFT.kk = \overline{EPSLii+RTDNii}$$

for $ii = A_E+1,\ldots,M+5$; where the terms involving retention rates vanish whenever the summation lower limit exceeds the upper limit. This is also true for $\Delta_s FTL..ii$ defined as follows:

for $N \geq 1$ and for $ii \leq \min\left[\pi(2)-\alpha+tt, M+5\right]$

$$\Delta_s FTL...ii = \sum_{jj=1}^{NCG34} \sum_{\substack{u=\pi(2) \\ kk=u-\alpha+1 \\ \gamma=ii-kk+1}}^{ii+\alpha-1} \Delta_s \eta(jj,u,\gamma) FTQCjjkk$$

$$+ \sum_{jj=1}^{NCG256} \sum_{\substack{u=\pi(2) \\ kk=u-\alpha+1 \\ \gamma=ii-kk+1}}^{ii+\alpha-1} \Delta_s \eta_2(jj,u,\gamma) FTQ2jjkk$$

Where:

for $N \geq 1$ and for $\pi(2)-\alpha+1+tt \geq ii \geq M+5$

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT-REQUIRED	PURPOSE

$$\Delta_s FTL...ii = \sum_{jj=1}^{NCG34} \sum_{\substack{u=ii+\alpha-tt \\ kk=u-\alpha+1 \\ \gamma=ii-kk+1}}^{u+\alpha-1} \Delta_s \eta(jj,u,\gamma) FTQCjjkk$$

$$+ \sum_{jj=1}^{NCG256} \sum_{\substack{u=ii+\alpha-tt \\ kk=u-\alpha+1 \\ \gamma=ii-kk+1}}^{ii+\alpha-1} \Delta_s \eta_2(jj,u,\gamma) FTQ2jjkk$$

$$+ \sum_{\substack{u=\pi(2) \\ kk=u-\alpha+1 \\ \gamma=ii-kk+1}}^{ii+\alpha-1-tt} \Delta_s \eta(u,\gamma) FTRtt.kk$$

<u>Note:</u> The terms involving the immediate reenlistment rates vanish when the immediate reenlistments are <u>not</u> computed by rates.

RIRQ in the first usage with FTQCjjkk represents IRG retention rates for FT with terms of service 3 and 4 years by month of reenlistment, by month of accession jj = C-group

RIR2 used with FTQ2jjkk, same as above except applies to FT with terms of service 2, 5, and 6 years.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - REQUIRED	PURPOSE
	RIRF used with FTRtt.kk represents IRG retention rates for FT enlistees of all terms of service, tracked as 1 C-group	
	RNSF Non-separation loss retention rates	
	where for $ii \leq \min[\pi(2)-\alpha+tt, M+5]$	
	$\Delta_s \eta(jj,u,\gamma)$ applies to FT with term of service 3 and 4 years where jj= C-group, u= month of accession, γ= length of service	
	$\Delta_s \eta_2(jj,u,\gamma)$ applies to FT with term of service 2, 5, and 6 years, where jj= C-group, u= month of accession, γ= length of service	
	for $\pi(2)-\alpha+1+tt \leq ii \leq M+5$	
	$\Delta_s \eta(u,\gamma)$ applies to FT, all terms of service	
E.ITRNii	$\text{ITRN..ii} - \sum_{\nu=-9}^{-1} f_1(\nu,ii) \vert \text{NPSG..kk}_{kk=ii+\nu} - \sum_{\nu=1}^{0} f_2(\nu,ii) \sum_{\substack{kk=\pi(2)-\alpha+1 \\ u=kk+\alpha-1}}^{ii+\nu-tt} \text{RIRF}(ii+\alpha-1+\nu,u)$	Computes transients; 12 of 36 accounts vary, the other 24 are summed intp the RHS by OTHRA.
	$\text{FTRtt.kk} - \sum_{\nu=1}^{0} f_2(\nu,ii) \sum_{\substack{kk=\pi(1)-\alpha+1 \\ u=kk+\alpha-1}}^{ii+\gamma-\text{MNSRIR}+1} \text{RIRD}(ii+\alpha-1+\nu,u)\omega\text{DRAFT.kk}$	

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - REQUIRED	PURPOSE

$$-f_3(ii)\text{ESEP..}ii - f_4(ii)\text{OFFL..}ii - \sum_{\nu=5}^{0} f_5(\nu,ii)\Big|_{kk=ii+\nu} \text{OFFG..}kk$$

$$-\sum_{jj=1}^{\text{NCG34}} \sum_{\nu=1}^{0} f_2(\nu,ii) \sum_{kk=\max(\pi(2)-\alpha+1,\,ii+\nu-tt+1)}^{ii+\nu-\text{MNSRIR}+1} \text{RIRQ}(ii+\alpha-1+\nu,u,jj)\text{FTQC}jjkk$$

for $ii \geq \pi(2)-\alpha+1+\text{MNSRIR}-1$

$$-\sum_{jj=1}^{\text{NCG256}} \sum_{\nu=1}^{0} f_2(\nu,ii) \sum_{kk=\max(\pi(2)-\alpha+1,\,ii+\nu-tt+1)}^{ii+\nu-\text{MNSRIR}+1} \text{RIR2}(ii+\alpha-1+\nu,u,jj)\text{FTQ2}jjkk$$

for $ii \geq \pi(2)-\alpha+1+\text{MNSRIR}-1$

$$= \text{OTHRA} + \sum_{\nu=1}^{0} f_2(\nu,ii)\left[\text{TIRG}(ii+\nu) + \sum_{\nu=9}^{0} f_6(\nu,ii)\text{REBR}(ii+\nu)\right]$$

for $ii=A_{TI}+1,\ldots,M+5$; (where $A_{TI} \geq A_E$ and where the terms involving retention rates RIRD and RIRF are not present when $ii+\nu<\pi(1)-\alpha+1+tt$ and $ii+\nu<\pi(2)-\alpha+1+tt$, respectively; or when the immediate reenlistments are user input. Any reenlistments not accounted for by the retention rates are included in the RHS).

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - REQUIRED	PURPOSE
	Notes: The index $ii+\alpha-1+\nu$ in RIRF, RIRD, RIRQ, and RIR2 gives the projection month (in the IPM scale) in which the immediate reenlistments occur; the second index u is the month of accession and in RIRQ nad RIR2, the third index represents the C-groups for terms of service 3, 4, and 2, 5, and 6 respectively.	
	jj=1,...,NCG34 jj=1,...,NCG256	
	RIRQ FT IRG retention rates, term of service 3 and 4	
	RIR2 FT IRG retention rates, term of service 2, 5, and 6	
	Terms involving the retention rates <u>vanish</u> whenever the summation lower limit exceeds the upper limit.	
<u>E.NPSGii</u>	$\text{NPSG..ii}-\omega\text{DRAFT.ii} - \sum_{jj=1}^{NCG34} \text{FTQCjjii} - \sum_{jj=1}^{NCG256} \text{FTQ2jjii} = 0$ for $ii=\min(A_{TE}+1,\ A_{TE}-3),\ldots,M+5$	Computes total NPS gains for each projection month.
<u>E.TRESii</u>	for $ii=A_{TE}+1,\ldots,M+5$ (where $A_{TE}\geq A_E$) $\text{TRES..ii}- \sum_{\nu=-8}^{0} f_1(\nu,ii)\ \eta(c,u,\gamma)\omega\text{DRAFT.kk}$ $c=1$ $u=\max(-\nu+1,ii+\alpha-1)+\nu$ $\gamma=-\nu+1$ $kk=ii+\nu$	Computes trainees.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - REQUIRED	PURPOSE

E.TRESii (cont.)

$$- \sum_{jj=1}^{NCG34} \sum_{\nu=-8}^{0} f_1(\nu,ii) \left| \eta(c,u,\gamma,jj) FTQCjjkk \right._{\substack{u=\max(-\nu+1,\,ii+\alpha-1)+\nu \\ \gamma=-\nu+1 \\ kk=ii+\nu \\ c=2}}$$

$$- \sum_{jj=1}^{NCG256} \sum_{\nu=-8}^{0} f_1(\nu\,ii) \left| \eta_2(c,u,\gamma,jj) FTQ2jjkk \right._{\substack{u=\max(-\nu+1,\,ii+\alpha-1)+\nu \\ \gamma=-\nu+1 \\ kk=ii+\nu \\ c=2}}$$

$$= \overline{TREADii} + \sum_{\nu=-8}^{0} f_2(\nu\,ii) \max_{jj} \left| \eta(c,u,\gamma,jj) PSG..kk \right._{\substack{c=2 \\ u=\max(-\nu+1,\,ii+\alpha-1)+\nu \\ \gamma=-\nu+1 \\ kk=ii+\nu}}$$

ETjj: $\eta(c=2,u,\gamma,jj)$,FT, term of service 3 and 4 years, characteristic group jj retention rates

E2jj: $\eta_2(c=2,u,\gamma,jj)$,FT, term of service 2, 5, and 6 years, characteristic group jj retention rates

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - REQUIRED	PURPOSE
L.FT..ii	$-\sum_{jj=1}^{NCG34} FTQCjjii - \sum_{jj=1}^{NCG256} FTQ2jjii + WNPSG.ii \leq 0$ for $ii=A_E+1,\ldots,M$	Constrains total FT enlistments in each month for all terms of service (2,3,4,5,6) to be at least as great as the WAC NPS gains in the same month.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
C.OBJ.04	$\sum_{jj=1}^{NCG34} \sum_{\substack{P=A_E+1 \\ ii=P}}^{M} [W_{QC}(jj,P)] [FTQCjjii] + \sum_{jj=1}^{NCG256} \sum_{\substack{P=A_E+1 \\ ii=P}}^{M} [W_{Q2}(jj,P)] [FTQ2jjii] \leq \bar{V}_4$	Constrains the value of the expression used for OBJ...04 to the neighborhood of the value obtained when OBJ...04 is used as the objective function. Prior to solution with OBJ...04, $\bar{V}_4$ is set equal to a large positive number. Subsequent to solution with OBJ...04, $\bar{V}_4$ is set to (value of the functional) + ε_4 where $\varepsilon_4 > 0$. Required if OBJ...04 is used.
EFTRttii	$FTRtt.ii - \sum_{\substack{jj=1 \\ R=ii \\ P=R+tt-1}}^{NCG34} [\eta(jj,R,P-R)] [FTQCjjii] - \sum_{\substack{jj=1 \\ R=ii \\ P=R+tt-1}}^{NCG256} [\eta_2(jj,R,P-R)] [FTQ2jjii] = 0$ for $ii=\pi(2)-\alpha+1,\ldots,M+5-tt$ if $N>1$ and $\pi(2)-\alpha \leq M+5-(tt+1)$	For each month for which COMPLIP determines the level of NPS gains in each C-group for terms of service 3 and 4, and 2,5, and 6, computes the total of these gains who remain in the Army at the end of the tt^{th} month of service. Required if $N>1$ and $\pi(2)-\alpha \leq M+5-(tt+1)$.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
E.FTL.ii	If N>1	Note: Only for N>1; see note on page K-4.

$$FTL...ii - \sum_{jj=i}^{NCG34} \sum_{\substack{R=\pi(2)-\alpha+1 \\ ii=P \\ kk=R}}^{P} [\lambda_E(jj,R,P)][FTQCjjkk]$$

$$-\sum_{jj=1}^{NCG256} \sum_{\substack{R=\pi(2)-\alpha+1 \\ ii=P \\ kk=R}}^{P} [\lambda_2(jj,R,P)][FTQ2jjkk] = \overline{LPFTii}$$

for $ii = P = \pi(2)-\alpha+1, \ldots, \min[\pi(2)-\alpha+tt,\ M+5]$,

If $\pi(2)-\alpha<M+5$;

$$FTL...ii - \sum_{jj=1}^{NCG34} \sum_{\substack{R=P-tt+1 \\ ii=P \\ kk=R}}^{P} [\lambda_E(jj,R,P)][FTQCjjkk]$$

$$-\sum_{jj=1}^{NCG256} \sum_{\substack{R=P-tt+1 \\ ii=P \\ kk=R}}^{P} [\lambda_2(jj,R,P)][FTQ2jjkk]$$

PURPOSE (continued): Computes losses for FT enlistments for terms of service 3 and 4 years and 2, 5, and 6 years for each month for which the number of enlistments is not user-specified to the IPM. This constraint is required if $\pi(2)-\alpha<M+5$, i.e., if FT enlistments of each C-group for each term of service are not user-specified inputs to the IPM for all projection months.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
	$- \sum_{\substack{R=\pi(2)-\alpha+1 \\ ii=P \\ kk=R}}^{P-tt} [\lambda_E(R,P)][FTRtt.kk] = \overline{LPFTii}$ for $ii = P = \pi(2)-\alpha+1+tt,\ldots,M+5$, if $\pi(2)-\alpha \leq M+5-(tt-1)$	
E.FQjjii	$FTQCjjii - FTQCjjkk = 0$ for $ii=\max[\pi(2)-\alpha+1, M+1],\ldots,M+5$ $kk=ii-12$ $jj=1,\ldots,NCG34$ IF $\pi(2)-\alpha<M+5$	Assigns values to FT enlistments, terms of service 3 and 4 years, of each C-group for five months beyond the end of the published manpower program Required if $\pi(2)-\alpha<M+5$.
EFQ2jjii	$FTQ2jjii - FTQ2jjkk = 0$ for $ii=\max[\pi(2)-\alpha+1, M+1],\ldots,M+5$ $kk=ii-12$ $jj=1,\ldots,NCG256$ IF $\pi(2)-\alpha<M+5$	Assigns values to FT enlistments terms of service 2,5, and 6 years of each C-group for five months beyond the end of the published manpower program. Required if $\pi(2)-\alpha<M+5$.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
L.AFQjji	$- \sum_{kk=\xi_i}^{\xi_{i+1}-1} FTQCjjkk + q_L(jj,i) \sum_{kk=\xi_i}^{\xi_{i+1}-1} (NPSG..kk-\omega DRAFT.kk) \leq 0$ For selected values of i in the range $1 \leq i \leq Y$ and any NCG34-1 or less values of jj in the range $1 \leq jj \leq NCG34$	Constrains FT enlistments, terms of service 3 and 4 years, C-group jj, to be no less than a specified fraction of total enlistments in selected FYs.
LAF2jji	$- \sum_{kk=\xi_i}^{\xi_{i+1}-1} FTQ2jjkk + q_2(jj,i) \sum_{kk=\xi_i}^{\xi_{i+1}-1} (NPSG..kk-\omega DRAFT.kk) \leq 0$ For selected values of i in the range $1 \leq i \leq Y$ and any NCG256-1 or less values of jj in the range $1 \leq jj \leq NCG256$	Constrains FT enlistments, terms of service 2, 5, and 6 years, C-group jj to be no less than a specified fraction of total enlistments in selected FYs.

Where in L.AFQjj $q_L(jj,i) \equiv$ lower limit on FT enlistments, C-group jj, terms of service 3 and 4 years, in year i specified as a fraction of total enlistments in that year, where jj and i have user-specified values.

LAFQ2jj $q_2(jj,i) \equiv$ lower limit on FT enlistments, C-group jj, terms of service 2, 5, and 6 years, in year i specified as a fraction of total enlistments in that year, where jj and i have user-specified values.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
LMFQjjii	$-\ FTQCjjii + \sum_{p=ii} [P_L(jj,P)]\ [NPSG..ii - \omega DRAFT..ii] \leq 0$ for selected values of ii in the range $A_E+1 \leq ii \leq M$ and any NCG34-1 or less values of jj in the range $1 \leq jj \leq NCG34$.	Constrains FT enlistments, TOS 3 and 4 years, C-group jj to be no less than a specified fraction of total enlistments in selected months.
LMF2jjii	$-\ FTQ2jjii + \sum_{p=ii} [P_2(jj,P)]\ [NPSG..ii - \omega DRAFT..ii] \leq 0$ for selected values of ii in the range $A_E+1 \leq ii \leq M$ and any NCG256-1 or less values of jj in the range $1 \leq jj \leq NCG256$.	Constrains FT enlistments, TOS 2, 5, and 6 years, C-group jj, to be no less than a specified fraction of total enlistments in selected months.

where $P_L(jj,P) \equiv$ Lower limit on FT enlistments, C-group jj, TOS 3 and 4 years, in month P, specified as a fraction of the total enlistments in that month where jj and P have user-specified values in the ranges $1 \leq jj \leq NCG34$, $A_E+1 \leq P \leq M$.

$P_2(jj,P) \equiv$ Lower limit on FT enlistments, C-group jj, TOS 2, 5, and 6 years, in month P, specified as a fraction of the total enlistments in that month where jj and P have user-specified values in the ranges $1 \leq jj \leq NCG256$, $A_E+1 \leq P \leq M$.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
LSFQjjii	$- FTQCjjii + \left[\prod_{p=ii}^{} s_L(jj,P)\right] \sum_{\substack{kk=\max(A_E+1,\xi_k) \\ P \epsilon \; FYk}}^{\xi_{k+1}-1} FTQCjjkk \le 0$ for selected values of jj, where $1 \le jj \le NCG34$, and ii, where $A_E+1<ii<M$. Note that the matrix coefficient for FTQCjjii is $[-1+s_L(jj,P)]$.	Imposes a lower limit on FT enlistments, TOS 3 and 4 years of C-group jj that reflects the seasonality of enlistments.
LSF2jjii	$- FTQ2jjii + \left[\prod_{P=ii}^{} s_2(jj,P)\right] \sum_{\substack{kk=\max(A_E+1,\xi_k) \\ P \epsilon \; FYk}}^{\xi_{k+1}-1} FTQ2jjkk \le 0$ for selected values of jj, where $1 \le jj \le NCG256$, and ii, where $A_E+1 <ii<M$. Note that the matrix coefficient for FTQ2jjii is $[-1+s_2(jj,P)]$.	Imposes a lower limit on FT enlistments, TOS 2, 5, and 6 years, of C-group jj that reflects the seasonality of enlistments.

Where $s_L(jj,P) \equiv$ Lower limit on FT enlistments of C-group jj, TOS 3 and 4 years, in month P, specified as a fraction of the total projected enlistments of C-group jj for the fiscal year in which the month P occurs, where jj and P have user-specified values in the ranges $1 \le jj \le NCG34$, $A_E+1 \le P \le M$.

$s_2(jj,P) \equiv$ Lower limit on FT enlistments of C-group jj, TOS 2, 5, and 6 years, in month P, specified as a fraction of total projected enlistments of C-group jj for the fiscal year in which the month P occurs, where jj and P have user-specified values in the ranges $1 \le jj \le NCG256$, $A_E+1 \le P \le M$.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
UAAFQjji	$\sum_{kk=\xi_i}^{\xi_{i+1}-1} FTQCjjkk \leq \overline{UAAFQjji}$ for selected values of jj, where $1 \leq jj \leq NCG34$ and i, where $1 \leq i \leq Y$.	Imposes an upper limit on the annual total FT enlistments for TOS 3 and 4 years of C-group jj to reflect the availability of recruits.
UAAF2jji	$\sum_{kk=\xi_i}^{\xi_{i+1}-1} FTQ2jjkk \leq \overline{UAAF2jji}$ for selected values of jj, where $1 \leq jj \leq NCG256$ and i, where $1 \leq i \leq Y$.	Imposes an upper limit on the annual total FT enlistments for TOS 2, 5, and 6 years of C-group jj to reflect the availability of recruits.
U.AFQjji	$\sum_{kk=\xi_i}^{\xi_{i+1}-1} FTQCjjkk - q_u(jj,i) \sum_{kk=\xi_i}^{\xi_{i+1}-1} (NPSG..kk - \omega DRAFT.kk) \leq 0$ for selected values of i in the range $1 \leq i \leq Y$ and any NCG34-1 or less values of jj in the range $1 \leq jj \leq NCG34$	Constrains FT enlistments, TOS 3 and 4 years, of C-group jj to be no more than a specified fraction of total enlistments in selected FYs.
U.AF2jji	$\sum_{kk=\xi_i}^{\xi_{i+1}-1} FTQ2jjkk - q_3(jj,i) \sum_{kk=\xi_i}^{\xi_{i+1}-1} (NPSG..kk - \omega DRAFT.kk) \leq 0$ for selected values of i in the range $1 \leq i \leq Y$ and any NCG256-1 or less values of jj in the range $1 \leq jj \leq NCG256$.	Constrains FT enlistments, TOS 2, 5, and 6 years, of C-group jj to be no more than a specified fraction of total enlistments in selected FYs.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
	Where $q_u(jj,i)$ ≡ Upper limit on FT enlistments, TOS 3 and 4 years, of C-group jj in year i, specified as a fraction of the total enlistments in that year, where jj and i have user-specified values in the ranges $1 \le jj \le NCG34$, $1 \le i \le Y$.	
	$q_3(jj,i)$ ≡ Upper limit on FT enlistments, TOS 2, 5, and 6 years of C-group jj in year i, specified as a fraction of the total enlistments in that year, where jj and i have user-specified values in the ranges $1 \le jj \le NCG256$, $1 \le i \le Y$.	
UMFQjjii	$FTQCjjii - \sum_{P=ii} P_u(jj,P)\left[NPSG..ii - \omega DRAFT.ii\right] \le 0$ for selected values of ii in the range $A_E+1 \le ii \le M$ and any NCG34-1 or less values of jj in the range $1 \le jj \le NCG34$.	Constrains FT enlistments, TOS 3 and 4 years of C-group jj to be no more than a specified fraction of total enlistments in selected months.
UMF2jjii	$FTQ2jjii - \sum_{P=ii} P_3(jj,P)\left[NPSG..ii - \omega DRAFT.ii\right] \le 0$ for selected values of ii in the range $A_E+1 \le ii \le M$ and any NCG256-1 or less values of jj in the range $1 \le jj \le NCG256$.	Constrains FT enlistments, TOS 2, 5, and 6 years of C-group jj to be no more than a specified fraction of total enlistments in selected months.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE
	Where $P_u(jj,P)$ ≡ Upper limit on FT enlistments, TOS 3 and 4 years of C-group jj in month P, specified as a fraction of the total enlistments in that month where jj and P have user-specified values in the ranges $1 \le jj \le NCG34$, $A_E+1 \le P \le M$.	
	$P_3(jj,P)$ ≡ Upper limit on FT enlistments, TOS 2, 5, and 6 years of C-group jj in month P, specified as a fraction of the total enlistments in that month where jj and P have user-specified values in the ranges $1 \le jj \le NCG256$, $A_E+1 \le P \le M$.	
USFQjjii	$FTQCjjii - \Big\vert_{P=ii} s_u(jj,P) \sum_{kk=\max(A_E+1,\xi_k)}^{\xi_{k+1}-1} FTQCjjkk \le 0$, $P \in FY_k$ for selected values or jj, where $1 \le jj \le NCG34$ and ii, where $A_E+1 \le ii \le M$. Note that the matrix coefficient for FTQCjjii is $[1-s_u(jj,P)]$.	Imposes an upper limit on FT enlistments, TOS 3 and 4 years, of C-group jj that reflects the seasonality of enlistments.
USF2jjii	$FTQ2jjii - \Big\vert_{P=ii} s_3(jj,P) \sum_{kk=\max(A_E+1,\xi_k)}^{\xi_{k+1}-1} FTQ2jjkk \le 0$ for selected values of jj, where $1 \le jj \le NCG256$ and ii, where $A_E+1 \le ii \le M$. Note that the matrix coefficient for FTQ2jjii is $[1-s_3(jj,P)]$.	Imposes an upper limit on FT enlistments, TOS 2, 5, and 6 years, of C-group jj that reflects the seasonality of enlistments.

COMPLIP-G3 CONSTRAINTS

ROW NAME	CONSTRAINT - OPTIONAL	PURPOSE

Where $s_u(jj,P)$ ≡ Upper limit on FT enlistments, TOS 3 and 4 years, of C-group jj in month P, specified as a fraction of the total projected enlistments of C-group jj for the fiscal year in which the month P occurs, where jj and P have user-specified values in the ranges $1 \leq jj \leq NCG34$, $A_E+1 \leq P \leq M$.

$s_3(jj,P)$ ≡ Upper limit on FT enlistments, TOS 2, 5, and 6 years, of C-group jj in month P, specified as a fraction of the total projected enlistments of C-group jj for the fiscal year in which the month P occurs, where jj and P have user-specified values in the ranges $1 \leq jj \leq NCG256$, $A_E+1 \leq P \leq M$.

Column name	Bounds - conditionally required
FTQCjjii	FT enlistments, TOS 3 and 4 years. Fixed for jj=1,...,NCG34 and $ii=\min(A_E+1,A_{TE}-7),\ldots,\pi(2)-\alpha$, if $\pi(2)-\alpha\geq\min(A_E+1,A_{TE}-7)$. Fixed for $ii=\xi_k\ldots,\min(A_E,A_{TE}-8),\ldots,A_E$ for any jj for which any of the following optional constraints are used: U.AFQjjk, L.AFQjjk, UAAFQjjk, for $k = \{\text{largest } k' \mid \xi_{k'} < \min(A_E+1,A_{TE}-7\}$.
FTQ2jjii	FT enlistments, TOS 2, 5 , and 6 years. Fixed for jj=1,...,MCG256 and $ii=\min(A_E+1,A_{TE}-7),\ldots,\pi(2)-\alpha$, if $\pi(2)-\alpha\geq\min(A_E+1,A_{TE}-7)$. Fixed for $ii=\xi_k,\ldots,\min(A_E,A_{TE}-8),\ldots,A_E$ for any jj for which any of the following optional constraints are used: U.AF2jjk, L.AF2jjk, UAAF2jjk, for $k = \{\text{largest } k' \mid \xi_{k'} < \min(A_E+1,A_{TE}-7)\}$.

APPENDIX I

MODIFICATIONS TO THE NON-PRIOR SERVICE GAINS (NPSG) MODULE

This appendix presents a brief resume of the ELIM-III Non-Prior Service Gains (NPSG) Module and indicates the enhancements undertaken to create an NPSG module compatible with ELIM-IV standards.

ELIM-III NPSG MODULE

The ELIM-III NPSG Module consists of the following components:

- NPSG Frequency File: the basic NPS accession information
- UPDATE: the NPSG Frequency File updating program
- PREPRO: the program interpreting and implementing user instructions
- BMD: the regression package creating accession projections
- HENRY: the graphical analysis program
- AGGRE: the C-group aggregation program

ELIM-IV MODIFICATIONS

NPSG Frequency File

The ELIM-III NPSG Frequency File is divided into two partitions. Partition 1 provides the following 192 disaggregations of male accessions.

- Age
  - Less than 18
  - 18, 19
  - 20, 21
  - Greater than 21
- Race
  - Black
  - Non-black

- Civilian education
  - Diploma high school graduate (DHSG)
  - General education development (GED)
  - Some high school
  - No high school
- Mental group
  - I, II, IIIA
  - IIIB
  - IV, V
- Enlistment bonus
  - Yes
  - No

Female accessions are assumed to be mental group I or II high school graduates (either diploma or GED) receiving no enlistment bonus and are disaggregated into four categories, as follows:

- Civilian education
  - DHSG
  - GED
- Race
  - Black
  - Non-black

Partition 2 of the ELIM-III frequency file disaggregates males by term of service into the following 94 categories:

- Age
  - All ages combined
- Race
  - Black
  - Non-black
- Civilian education
  - DHSG
  - GED
  - Not DHSG or GED

- Mental group
  - I, II, IIIA
  - IIIB
  - IV, V
- Term of enlistment
  - 2 years
  - 3 years
  - 4, 5, 6 years

The ELIM-IV NPSG Frequency File eliminates the two-partition concept and disaggregates accessions into 768 categories, which are all combinations of levels of the following factors:

- Age
  - Less than 18
  - 18, 19
  - 20, 21
  - Greater than 21
- Race
  - Black
  - Non-black
- Civilian education
  - DHSG
  - Not DHSG (NHS)
- Mental group
  - I, II, IIIA
  - IIIB
  - IV, V
- Sex
  - Male
  - Female
- Term of enlistment
  - 2 years
  - 3 years
  - 4 years
  - 5, 6 years

- Enlistment bonus
  - Yes
  - No

The ELIM-IV NPSG Frequency File is an IRIS-compatible direct access file configured to contain 84 months of data beginning with January 1972. The ELIM-III Frequency File was a sequentially organized disk file capable of containing 72 months of data beginning with February 1972.

UPDATE: The NPSG Frequency File Updating Program

Due to the wholesale restructuring of the NPSG Frequency File, a completely new version of UPDATE has been written to maintain the ELIM-IV file.

PREPRO: User Instructions

Under the ELIM-III system, aggregation of individual population cells to form dependent variables is accomplished by listing the cell numbers to be aggregated for a particular variable on input cards. This system is very tedious for the 196 cells used in the ELIM-III Partition 1 but altogether too cumbersome for the 768 cells used in the ELIM-IV data file.

Because of this unwieldy input requirement and a desire for strict editing of user specifications, a completely prompted interactive front-end program (QUERY) was developed. QUERY permits dependent variables to be defined by entering specifications as indicated in Exhibit I.1.

Additionally, all other information required for the execution of the ELIM-IV NPSG Module is prompted for, edited, and recorded for use by the component programs of the module.

The ELIM-III version of PREPRO has been extensively modified to accept as input the card-image file created by QUERY rather than a user-created card deck.

BMD: The Regression Package

The BMD regression package is essentially unchanged. With the exception of several small improvements in efficiency, the only enhancement provided in the ELIM-IV BMD is the capability of producing forecasts for as many as 89 months past the end of historical data.

HENRY: The Graphical Analysis Program

The graphical analysis program is unchanged except that forecasts for months which cannot be accommodated on the maximum graph width are not graphed at all.

AGGRE: The C-Group Aggregation Program

The C-group aggregation program is unchanged except that C-group aggregations can now be provided to the IPM as well as to COMPLIP through the Matrix Generator. The maximum number of C-groups which can be passed to the IPM or the Matrix Generator has been increased to 40 to satisfy the requirements of the C-group expansion task.

Exhibit I.1

```
SPECIFY DESIRED AGGREGATIONS BY ENTERING IN THE INDICATED FIELDS
   (1) VARIABLE NAME AND
   (2) CODES INDICATING CHARACTERISTICS TO BE COMBINED (LEFT-JUSTIFIED).
MULTIPLE CODES FOR A CHARACTERISTIC ARE ENTERED IN A SERIES WITHOUT
   DELIMITERS.
```

CHARACTERISTIC	CODE	DEFINITION
AGE	1	LESS THAN 18
	2	18,19
	3	20,21
	4	MORE THAN 21
	9	ALL OF THE ABOVE
RACE	1	BLACK
	2	NON-BLACK
	9	ALL OF THE ABOVE
EDUCATION	1	DHSG
	2	NHS,GED
	9	ALL OF THE ABOVE
MENTAL GROUP	1	I,II,IIIA
	2	IIIB
	3	IV
	9	ALL OF THE ABOVE
SEX	1	MALE
	2	FEMALE
	9	ALL OF THE ABOVE
TERM	1	5,6 YEARS
	2	2 YEARS
	3	3 YEARS
	4	4 YEARS
	9	ALL OF THE ABOVE
BONUS	1	YES
	2	NO
	9	ALL OF THE ABOVE

ENTRY OF "*FORMAT" FOR "NAME" WILL REPRODUCE FORMATTING HEADER

		CODES						
VAR	NAME	AGE	RACE	ED	MG	SEX	TERM	BONUS
1	FEMALES	9	9	9	9	2	9	9
2								