

AD-A194 044

THE RESERVE COMPONENTS COMMON PERSONNEL DATA SYSTEM
(RCCPDS) FOR NAVAL RE. (U) CENTER FOR NAVAL ANALYSES
ALEXANDRIA VA NAVAL PLANNING MANPOW. M C REGETS

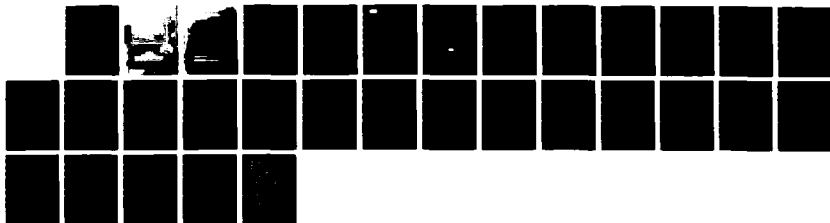
1/1

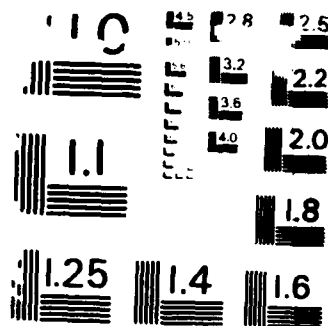
UNCLASSIFIED

JUN 87 CRM-87-106 N00014-87-C-0001

F/G 15/1

NL

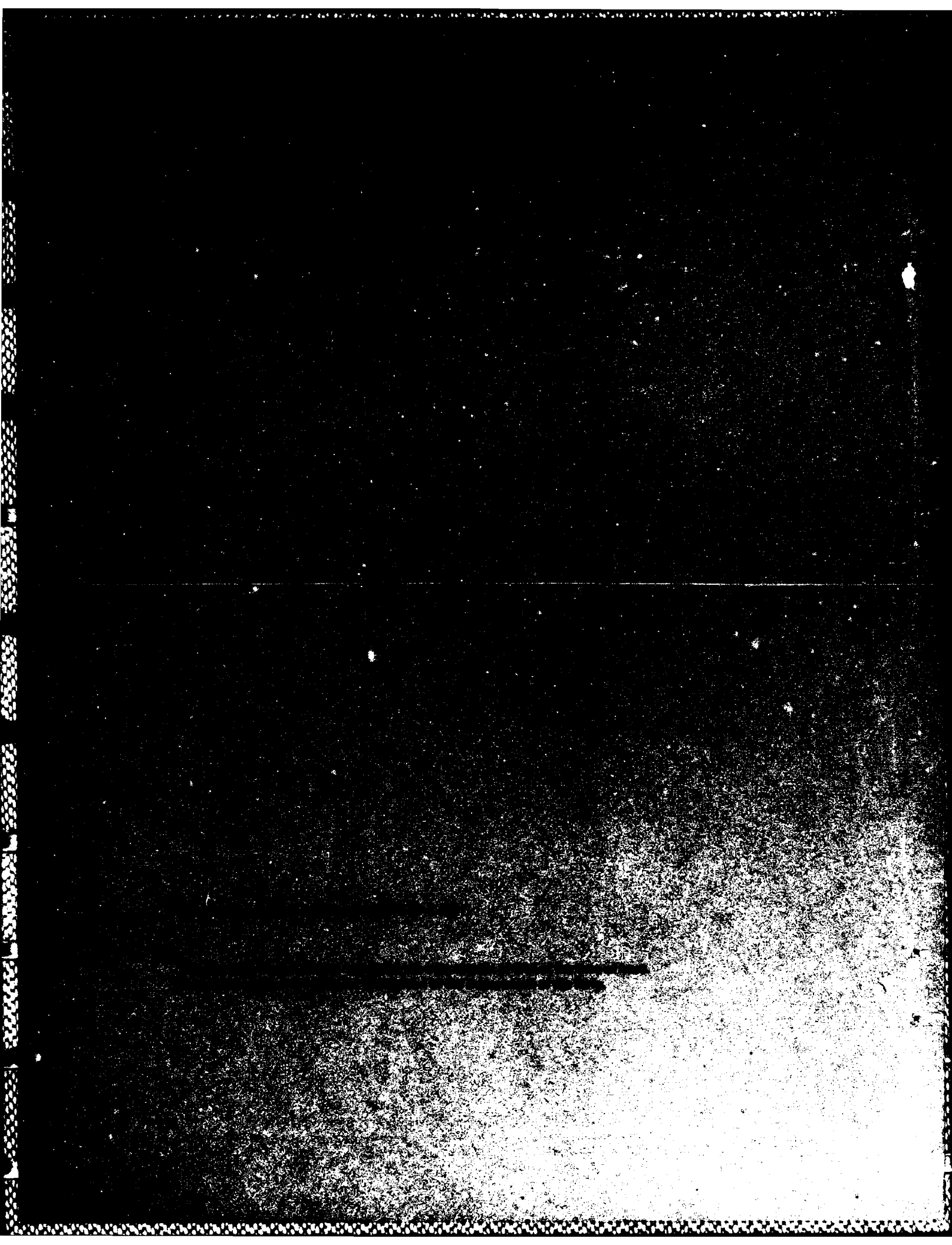




COPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A194 044

0 0 0 0 0 0



REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION / AVAILABILITY OF REPORT		
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE			APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CRM 87-106			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION Center for Naval Analyses		6b. OFFICE SYMBOL (if applicable) CNA	7a. NAME OF MONITORING ORGANIZATION Office of Chief of Naval Operations (OP-09R)		
6c. ADDRESS (City, State, and ZIP Code) 4401 Ford Avenue Alexandria, Virginia 22302-0268			7b. ADDRESS (City, State, and ZIP Code) Navy Department Washington, D.C. 20350-2000		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research		8b. OFFICE SYMBOL (if applicable) ONR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-87-C-0001		
8c. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, Virginia 22217			10. SOURCE OF FUNDING NUMBERS		
			PROGRAM ELEMENT NO. 65154N	PROJECT NO. R0148	TASK NO.
			WORK UNIT ACCESSION NO.		
11. TITLE (Include Security Classification) Reserve Components Common Personnel Data System (RCCPDS) for Naval Reserve Enlisted Personnel					
12. PERSONAL AUTHOR(S) Mark C. Regets					
13a. TYPE OF REPORT Final		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Year, Month, Day) June 1987	
15. PAGE COUNT 26					
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
05	02				
05	09				
19. ABSTRACT (Continue on reverse if necessary and identify by block number) This paper discusses the Reserve Components Common Personnel Data System (RCCPDS), one important source of Reserve manpower data. The RCCPDS transaction file provides data on gains, losses, and reenlistments for each military service's Selected Reserve (SELRES). The strengths and weaknesses of the RCCPDS transaction file as a source of data on Navy SELRES manpower are examined, with particular emphasis on enlisted personnel.					
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified		
22a. NAME OF RESPONSIBLE INDIVIDUAL Cdr. Merritt			22b. TELEPHONE (Include Area Code) 657-5201		22c. OFFICE SYMBOL OP-09R5

CRM 87-106

18. System), SELRES (Selected Reserve), Statistical data



CENTER FOR NAVAL ANALYSES

A Division of Hudson Institute 4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

7 Jul 1987

MEMORANDUM FOR DISTRIBUTION

Subj: Center for Naval Analyses Research Memorandum 87-106

Encl: (1) CNA Research Memorandum 87-106, "The Reserve Components Common Personnel Data System (RCCPDS) for Naval Reserve Enlisted Personnel," by Mark C. Regets, Jun 1987

1. Enclosure (1) is forwarded as a matter of possible interest.
2. This Research Memorandum discusses the Reserve Components Common Personnel Data System (RCCPDS), one important source of Reserve manpower data. The RCCPDS transaction file provides data on gains, losses, and reenlistments for each military service's Selected Reserve (SELRES). The strengths and weaknesses of the RCCPDS transaction file as a source of data on Navy SELRES manpower are examined, with particular emphasis on enlisted personnel.

Robert F. Lockman

ROBERT F. LOCKMAN
Director
Manpower Program

Distribution list:

CAPT Hetherington (OP-09R5)
CAPT Glad (OP-09R9)

CRM 87-106 / June 1987

THE RESERVE COMPONENTS COMMON PERSONNEL DATA SYSTEM (RCCPDS) FOR NAVAL RESERVE ENLISTED PERSONNEL

Mark C. Regets

Naval Planning, Manpower, and Logistics Division

A Division of

CNA

Hudson Institute

CENTER FOR NAVAL ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

This paper discusses the Reserve Components Common Personnel Data System (RCCPDS), one important source of Reserve manpower data. The RCCPDS transaction file provides data on gains, losses, and reenlistments for each military service's Selected Reserve (SELRES). The strengths and weaknesses of the RCCPDS transaction file as a source of data on Navy SELRES manpower are examined, with particular emphasis on enlisted personnel.



Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

TABLE OF CONTENTS

	<u>Page</u>
Introduction	1
Background	1
Origin of RCCPDS Data	2
RCCPDS Versus the IEMF	3
RCCPDS Transaction File: Data Characteristics and Problems	6
RCCPDS Data Elements	6
Accession Program Identification	8
RCCPDS Transaction Patterns	9
Illogical Gain/Loss Patterns	9
Problems With Reenlistment Transactions	11
Different Transactions on the Same Date	11
The Age of Transactions	12
Conclusion	13
Appendix A: Options for RCCPDS Data Cleanup	A-1
Appendix B: RCCPDS Data Formats	B-1 - B-2
Appendix C: Notes on RCCPDS Data Elements	C-1

INTRODUCTION

This research memorandum discusses the Reserve Components Common Personnel Data System (RCCPDS), one important source of Reserve manpower data. The RCCPDS transaction file provides data on gains, losses, and reenlistments for each military service's Selected Reserve (SELRES). The strengths and weaknesses of the RCCPDS transaction file as a source of data on Navy SELRES manpower are examined, with particular emphasis on enlisted personnel.

Little historic manpower data are available at the individual level for the U.S. Naval Reserve. Although separate data files are created by many different institutional data users, these files have not been systematically preserved. The files are intended to serve specific management functions rather than to provide historic databases for analysis.

The historic files that do exist often have missing or inconsistent data. Quality-control efforts by the managers of these databases have focused on data elements needed for the day-to-day management of the Naval Reserve. The data needed to support medium- and long-range planning often receive much less attention. For example, the date when an individual entered the Selected Reserve was not recorded on any computerized personnel file until October 1986. Demographic data not affecting pay, such as age and education, are often missing or have questionable values. Even pay-related data, such as length of service, are not updated for all individuals on all files.

The value of the RCCPDS transaction file lies in its preservation of past Naval Reserve data that the Defense Manpower Data Center (DMDC) uses to create the file. The RCCPDS provides information on Navy SELRES gains and losses for a longer period than is otherwise available. In addition, it captures information on individuals at the time of their SELRES affiliation. However, some data elements are lost or recoded over the course of an individual's SELRES service.

The RCCPDS also has several different types of data problems. Many are caused by errors and limitations in the original Naval Reserve data. Other problems result from the creation of a transaction file (containing individual gains and losses) from manpower inventory data. Some other limitations come simply from conversion of data into standard Department of Defense (DOD) codes. This paper discusses each problem in detail. Possible ways to deal with RCCPDS data problems are discussed in appendix A.

BACKGROUND

The RCCPDS is a source of data for both officer and enlisted personnel in the Reserves of all U.S. military services. RCCPDS is maintained by the DMDC, which receives data monthly from each of the

seven Reserve components. The Coast Guard Reserve, National Guard, and Air National Guard are included along with the reserve components of the four major services. Separate computer programs, with decision rules adapted for each service's different personnel systems and data, translate the monthly data submissions into standard DOD descriptions for RCCPDS.

There are two types of RCCPDS files--a master file and a transaction file. The master file contains monthly information on all personnel belonging to any category of Reserve Force, including the Standby and Retired Reserves. The transaction file reports gains, losses, and reenlistments for only the SELRES. Each record in the transaction file contains the full master-file record for the individual at the time of the transaction, and indicates the type of transaction and the date when it occurred. The transaction date is of particular importance for Naval Reserve studies because the initial SELRES affiliation date is not otherwise recorded in RCCPDS or other data sources.

Origin of RCCPDS Data

RCCPDS data for Navy enlisted personnel are derived from the Naval Reserve Personnel Center's (NRPC) Inactive Enlisted Master File (IEMF). The IEMF is an enlisted personnel inventory file for all parts of the Naval Reserve, including SELRES, Individual Ready, Fleet, and Retired Reserves. Navy Officer records in RCCPDS, not examined in detail here, come from a similar Inactive Officer Master File.

IEMF data come from two principal sources. Some active-duty data are transferred by NRPC from the Navy active-duty Enlisted Master Record. For SELRES personnel, most data originate with weekly reports received from each SELRES unit. These reports are made on special optical character recognition forms that can be easily transferred to a computer. The reports are commonly known as Reserve Field Reporting System (RESFIRST) diaries. RESFIRST diaries are the usual source of all updates and corrections to the IEMF.

Each month's IEMF is compared to the previous month to extract the data needed to generate a submission file for RCCPDS transactions. This cannot be a simple check of the presence of an SSN on the file, since the IEMF is an inventory file for all Naval Reserve categories, not just SELRES. Changes in the IEMF data-elements strength code (SC), training category (TCAT), type loss (TYL), and Reserve Forces Category (RFC) are all examined to determine if a SELRES transaction has taken place. A series of decision rules is necessary to create RCCPDS transactions, since there is no one place on the IEMF that signals the occurrence and type of transaction.

None of these IEMF data elements are exclusively for SELRES personnel. RFC identifies SELRES, Individual Ready Reserve (IRR), and

all other categories of Reservists. TCAT identifies the type of SELRES drill obligation and the Stand-by Reservists and SELRES personnel in or awaiting training. Strength code indicates whether an individual is counted towards manpower strength at an activity. Type loss refers to a loss from the IEMF, not to a type of SELRES loss. Type loss contains information on a SELRES loss only if the individual is not transferred to the IRR.

When there is a logical conflict between data elements, a second level of decision rules is used in the RCCPDS creation program to try to resolve the conflict. For example, a blank RFC is a valid code for the Retired Reserve, but sometimes occurs with a TCAT code of A, indicating a drilling member of SELRES. RCCPDS prior to October 1986 would consider that individual a member of SELRES, but in the latest revision would not.

The NRPC sends the RCCPDS transaction-submission file to the DMDC. A transaction is accepted if it is consistent with the previous month's RCCPDS master file. Thus, a gain transaction would not be recorded if the individual already appears on the previous month's master file as part of the SELRES inventory. Similarly, a loss will not be accepted unless an individual appears in SELRES in the previous month or has a gain transaction in that same month.

Generally, the cleaning performed by DMDC will create new data problems only if the reporting of an individual's SELRES loss is processed before his gain or if a reaffiliation is processed before the preceding loss. In such a case, valid transactions may be deleted, leaving a seemingly illogical transaction pattern.

RCCPDS Versus the IEMF

It is usually better to use data from as original a source as possible. The IEMF is the sole source of data for RCCPDS on Naval Reserve enlisted personnel. However, several practical limitations make the IEMF a difficult data set to use for many analytic purposes.

The IEMF is a personnel-inventory file, not a transactions file of gain and loss records. Furthermore, it is an inventory file for all Naval Reserve programs, not just the Selected Reserve. This makes it difficult to determine the date of individual SELRES gains and losses. The date of an individual's gain to the Selected Reserve was not recorded until October 1986. Thus, SELRES gains and losses must be inferred from changes in IEMF data fields. Such inferences require a set of decision rules and computer matches of successive months of the IEMF, a process similar to the creation of the RCCPDS transaction file.

Because monthly IEMF historic tapes are unavailable, gain/loss records can be created by matching the files for only a few years of data. Quarterly IEMF files are available at CNA from June 1985 to the

present. September IEMFs are available from 1982, and September extracts of SELRES drillers from 1976. NRPC does not keep historical data, and other repositories of the IEMF have not been found. It is thus impossible to go back to the IEMF, the source of the RCCPDS transactions, and use different decision rules to generate transactions.

Many of the IEMF data elements used by RCCPDS are subject to change over time. Information on gains and losses to the IEMF (not necessarily a gain or loss to SELRES) is kept on the file for only three months. Many other data elements, such as the date the individual was received by a unit (DRCDD), may be initially correct, but suffer from mistakes made in subsequent RESFIRST diary entries submitted by SELRES units. This is a particularly serious problem because the IEMF has no data element that provides the date an individual entered SELRES. When the IEMF record is examined at the time the transaction first appears, both the date and the existence of the transaction can be more accurately identified. This is exactly what is attempted when the RCCPDS transaction file is created.

The date of a transaction, determined from the DRCDD, can be different from the date when it first appears on the IEMF. The date when a transaction is added to RCCPDS is known as its tapedate. The age of a newly reported transaction is measured as the difference between the transaction's tapedate and its effective date.

There can be long lags before the reporting and/or processing of gains and losses. Lags in reporting losses are slightly longer than for gains. Part of the delay for losses could come from the requirement that an individual miss five drills (usually over a two-month period) before a unit can remove him for poor attendance.

The problem with lagged losses does persist for more than just two months. Figure 1 shows the distribution of reporting dates relative to effective dates for gains and losses on the FY 1981 through FY 1985 RCCPDS transaction files, using only effective dates from the beginning of FY 1981 through the end of FY 1984. The RCCPDS program for transaction creation obtains the effective date of a transaction from one of two DRCDD fields on the IEMF at the time the transaction is generated. A comparison of the effective date and the date when RCCPDS created the transaction shows the length of any delay in reporting.

The limited availability of IEMF data can create a number of biases when SELRES affiliation and continuation are estimated using only the available IEMF data. The use of once-a-year "snapshots" of the IEMF SELRES inventory tend to bias estimates in the following ways:

- Many personnel with short continuation leave SELRES without being observed in the snapshot. This reduces the number of observed affiliations, and over-samples

individuals with better continuation. In addition, continuation estimates are biased upwards.

- The inability to observe the exact date of affiliation also biases continuation estimates upwards. All anecdotal and empirical evidence suggests that month to month continuation dramatically improves the longer the individual remains in SELRES. Thus, the percentage of personnel surviving the first year of SELRES should be smaller than the percentage surviving from the seventh through the eighteenth month. A once-a-year inventory file cannot distinguish between these different periods of service.
- Lags in data reporting also affect both affiliation and continuation estimates. A lag in reporting a gain reduces the chance that an affiliation will be observed within a given period from the end of an individual's active duty. Longer lags for losses, a less serious problem, also lead to over-estimates for continuation.

These biases all operate in the same direction, thereby magnifying the effect of data errors in the IEMF.

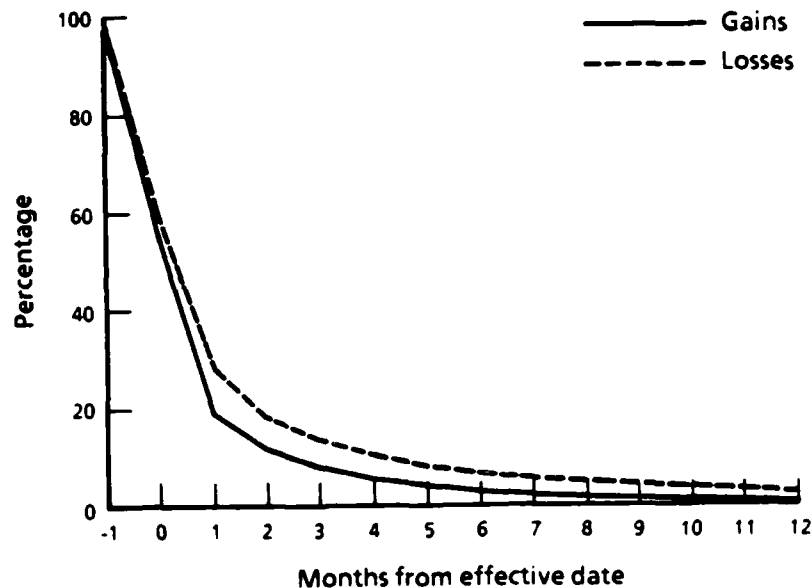


FIG. 1: PERCENTAGE OF OUTSTANDING TRANSACTIONS

RCCPDS TRANSACTION FILE: DATA CHARACTERISTICS AND PROBLEMS

This section discusses some of the more important characteristics of the RCCPDS Transaction File for Naval Reserve enlisted personnel. The development of new data elements for RCCPDS, as well as the creation of transaction records from an inventory file, creates both advantages and problems for analysts.

The advantages include the easy accessibility of SELRES gain and loss information and, for some purposes, the standardization of data codes among military services. As a practical matter, the latter advantage is useful only to those doing cross-service studies with access to RCCPDS files for all military services.

Most problems found in transaction-file data can be placed in one of two categories: missing or miscoded data, and inconsistent or unlikely patterns of transactions for individuals. Miscoded or missing data create a number of problems, including making it difficult to identify SELRES accession programs and drill obligations. Problems with the patterns of individual transactions, such as an individual with two gains without an intervening loss, suggest that some transactions listed in RCCPDS may be erroneous. Unusual transaction patterns for an individual can include illogical patterns of gains and losses, more than one transaction on the same date, and reenlistment transactions inconsistent with the individual's gain or loss records.

RCCPDS Data Elements

As a standardized DOD data file, RCCPDS does suffer from the generalization necessary to accommodate data from all services. This sometimes leads to excluded data, missing data, and attempts to code RCCPDS data elements when they cannot be adequately determined from Naval Reserve data.

The most useful data elements are the transaction effective date and the transaction code. It is these data that are not easily available from other sources. A transaction code beginning with 'G' indicates a gain, which would make the transaction effective date the date of the individual's gain to SELRES. When the transaction code begins with 'L,' the date of the individual's SELRES loss is given. When this information is matched using social security numbers, the length of an individual's service in SELRES can be determined.

Table 1 shows the data elements for enlisted personnel contained in the RCCPDS transaction file. Formats and codes for each data element can be found in DOD Instruction 7730.54.¹ The percentage of each data

1. Department of Defense Instruction Number 7730.54, "Reserve Component Common Personnel Data System," 26 Oct 1981.

TABLE 1

RCCPDS DATA ELEMENTS: NAVY ENLISTED PERSONNEL

<u>Data element</u>	<u>Percentage missing</u>
Reserve component	N.A.
Social Security number	0.0
Name	0.0
Reserve component category	0.0
Pay grade	0.2
Date of pay grade	0.2
Full time support status	83.2
Active Guard/Reserve indicator	83.4
Date of birth	0.9
End of term of service (ETS)	3.2
State of residence	1.5
Zip code	1.6
Pay entry base date	0.1
Rating	0.0
Duty rating	66.7
Source of first entry	2.4
Race	5.0
Ethnic group	33.5
Sex	0.0
Number of dependents	0.0
Civilian education	0.2
Professional military education	100.0
Language	100.0
Total active federal military service	32.5
Years of service for retirement	99.8
Program element code	40.0
Unit identity code	21.5
Unit state	18.2
Unit zip code	20.4
Basic branch	100.0
Current aeronautical rating	98.4
Current flying status	99.9
Eligible for 20 year retirement	97.0
Incentive status date	98.4
Incentive status code	98.3
Date of initial military service	100.0
Date of initial reserve service	100.0
Civilian occupation code	100.0
Mental group	38.3
Obligor status	87.3
Term of enlistment	22.4
Disputed data indicator	100.0
Transaction code	0.0
Separation program designator	89.7
Character of service	89.7
Transaction effective date	0.0
Tape date	0.0

element missing was calculated from a sample of 12,000 records from the FY 1981 through FY 1986 files for enlisted personnel. A large proportion of missing values does not necessarily mean that the data element is incomplete, since a blank is sometimes used instead of a zero as a valid code. For example, Obligor Status appears reasonable with 87.3 percent missing because most Navy SELRES personnel do not have a SELRES obligation. However, the 98.4 percent missing for Incentive Status Code is too large to represent the portion of personnel ineligible for bonuses.

Some data elements of particular interest to the Navy are excluded. For example, an individual's Navy SELRES accession program can only be inferred from the RCCPDS code for type of gain transaction. Another example is the data field that holds a military occupational specialty (MOS) code for other services. For the Navy, this contains the numeric and letter codes for Naval Reserve enlisted ratings, but not an individual's Naval Enlisted Classification (NEC). NECs would be necessary to provide occupation descriptions that are as detailed as those provided by MOS. For example, the Army has a MOS code for flute player, while the Navy rating identifies such an individual only as a musician.

Other data elements of analytic interest are given fields in RCCPDS, but are always missing for Navy personnel. For example, the dates for initial military and initial Reserve service are missing because that information is not available on the IEMF. Although incentive-pay status in RCCPDS is intended to reflect bonuses for affiliation or reaffiliation in SELRES, the IEMF contains information only on incentive pay for special duties.

Other comments and notes on RCCPDS data elements are included in appendix C.

Accession Program Identification

The identification of an individual's SELRES accession program using only the RCCPDS transaction file is inexact in many ways. For example, recruits from other military services (OSVETs) who do not administratively pass through their original service's IRR cannot be distinguished from Navy veterans. Advanced Pay Grade recruits (APGs) who have prior military service are correctly identified as veterans, but not as members of the advanced pay-grade program. If possible, program identification should be done with the assistance of other data sources.

Rough identification of SELRES accession programs can be made in a number of ways. Non-prior-service (NPS) recruits should enter SELRES with a transaction code of 'G1' and a Source code of '6,' both directly indicating NPS status. Instead, a large number of Naval Reserve NPS recruits have codes of 'G3' and '2,' indicating a Navy veteran entering

SELRES from civilian life. These can be distinguished by using other data fields, which show zero to six months of total active military service and pay grades of E1 or E2. The Active Mariner program is identified by a Source Code of '5' and a transaction code of 'G3,' 'G4,' 'G6,' or 'G9,' which indicate a veteran of the same military service. Some OSVETs can be identified by a transaction code of 'G5,' indicating a transfer from the reserve of another service. Non-prior-service APGs should have pay grades at the time of their gain of E4 or E5.

RCCPDS Transaction Patterns

The RCCPDS transaction file, as discussed above, is created from Naval Reserve manpower-inventory files. A decision to create an RCCPDS gain, loss, or reenlistment is made through a comparison of two successive monthly IEMFs. A transaction based on the application of succession rules to IEMF data will be rejected by DMDC only if it is inconsistent with the inventory file from the prior month. No checks are made of previous RCCPDS transaction files.

It is thus possible to have patterns of RCCPDS transactions that are not logically consistent. A gain on one date can be followed by another gain without a report of an intervening loss. Double losses are also observed. These patterns occur both as duplicate records with identical effective dates and with different effective dates. Reenlistments can also have effective dates that make them inconsistent with other transactions.

Of the 192,951 individuals on the Navy enlisted RCCPDS transaction files for FY 1981 through FY 1985, 2.1 percent had illogical gain/loss patterns, 2.0 percent had reenlistment code problems, and 4.0 percent had more than one transaction on the same day. The proportion of individuals exhibiting any of the three data conditions is 6.0 percent. There are overlaps between each category. A further breakdown of each category is given in table 2.

Illogical Gain/Loss Patterns

An illogical gain/loss pattern is defined as two gains without an intervening loss or two losses without an intervening gain for the same individual. This is determined from the sequence of effective dates for the transactions. Illogical gain/loss patterns occurred for 2.1 percent of individuals on the FY 1981 through FY 1985 RCCPDS transaction file.

The occurrence of an illogical gain/loss pattern does not seem to be related to any of the major individual characteristics reported on RCCPDS. An individual's SELRES accession program does not affect the chance of having an illogical pattern. There are no significant differences by pay grade, rating group, age, race, or sex. It is not clear what this implies about how erroneous transactions are generated.

TABLE 2

PERCENTAGE OF PROBLEMS IN INDIVIDUAL PATTERNS
(192,951 Individual Transaction Patterns)

Any unusual pattern	6.0
Illogical gain/loss	2.1
Reenlistment code problem	2.0
Transactions on same date	4.0
Illogical gain/loss	2.1
Double gain or loss--same date	0.4
Other with transactions on same date	0.5
Illogical, different dates	1.7
Transactions on the same date	4.0
Gain and loss	3.0
Duplicate transactions	0.5
Reenlistment with gain or loss	0.4
Reenlistment code problem	2.0
Duplicate reenlistment	0.4
Reenlistment after a loss	1.5
Reenlistment before a gain	0.4
With gain or loss on same date	0.4

Individuals with illogical patterns do have statistically significant differences in their duration in SELRES. The observed average SELRES duration for individuals with no illogical pattern is 12.1 months. Taking the time from the first recorded gain, observed SELRES duration is 9.2 months for those with double gains or double losses. Due to the small portion of total patterns with these problems, inclusion of the illogical data would lower duration estimates by only 0.1 month.

Many individuals with illogical patterns had transactions with the same effective date. In 0.4 percent of the total patterns, two gain or two loss transactions had the same effective date. Thus, a gain or a loss will appear on RCCPDS one month and appear again in a later month with the same effective date. When the second transaction is just a duplicate of the first, there are no analytic problems posed so long as there is no reason to believe that both transactions were generated erroneously. There is no conflict concerning the date of occurrence for the two transactions. In another 0.5 percent of individual patterns, an illogical gain/loss pattern results when one of the conflicting transactions shares an effective date with a third transaction.

Problems With Reenlistment Transactions

Reenlistments and extensions in SELRES are also coded as RCCPDS transactions and can create or be part of illogical transaction patterns. A reenlistment record occurs after a loss without an intervening gain for 1.5 percent of individuals. More commonly, a reaffiliation to SELRES after a lapse in service is reported as a gain.

A reenlistment transaction before a gain occurs for 0.4 percent of individuals. It is not clear when SELRES service began for these individuals. If the reenlistment transaction should have been recorded as a gain, why is there a later gain?

Different Transactions on the Same Date

The existence of two transactions for the same individual with the same date is not necessarily a problem, but is a cause for concern. A simultaneous gain and loss occurs for 3.0 percent of individuals on the FY 1981 through 1985 files. A simultaneous reenlistment and gain or reenlistment and loss occurs for another 0.4 percent. Because many Navy veterans do stay in SELRES for only one month, both their gain and loss dates could be reported as the date of their only drill. Some of these simultaneous gains and losses, however, occurred for individuals with a preceding gain record. If an individual is lost and regained on the same day, this should be properly coded as a reenlistment.

A separate case occurs when there are three or more transactions with the same effective date. Except when the pattern is triple gain or triple loss, it is not obvious whether a gain or loss transaction, or both, should be recorded. These occur for 0.8 percent of individuals, accounting for 26.1 percent of simultaneous gain/loss patterns.

The interpretation is also unclear when a reenlistment code occurs as part of a simultaneous pattern. Did an individual stay or go when a reenlistment and loss transaction have the same effective date? This is observed for 0.2 percent of enlisted personnel in the FY 1981 through FY 1985 files. A simultaneous gain and reenlistment occurs for 0.1 percent of these individuals.

Multiple transactions, created on the same month by errors in the RCCPDS creation program, do not appear to be a significant problem. The binary coded field on RCCPDS, known as tapedate, reports the year and month when a transaction has been added to the file. Only 149 out of 192,951 individuals on the FY 1981 through FY 1985 files have any duplicate tapedate. The 4.0 percent that have transactions occurring on the same effective date are almost all reported to RCCPDS on a different month.

A test was also performed using the RCCPDS creation program and CNA's copies of the August and September 1984 IEMF. No more than one

transaction was generated for any one SSN. Many of the possible problems with the program would have resulted in more than one transaction generated for the same individual in the same month. This does not rule out the possibility that some transactions were erroneously generated, but does indicate that the decision rules of the RCCPDS COBOL program are less likely to cause problems.

There are many other hypothetical stories describing how illogical transaction patterns are generated, but few can be empirically tested. For example, a second gain could be the miscoding of a reenlistment, but data on contract length are not good enough to verify this. There is inconsistency between gain date, contract length, and end of term of service for 54.2 percent of the logical patterns. Similarly, Pay Entry Base Date (PEBD) should change whenever there is a gap in an individual's SELRES or active service. PEBD did not usually change on either the IEMF or RCCPDS even when both data sources indicate a lapse in service.

The Age of Transactions

Since RCCPDS transactions are created from the IEMF, the time lags in reporting events affect RCCPDS as well. Although its reported effective date should be correct, a transaction may not appear on RCCPDS for some time. This creates a practical limitation on the use of recently generated RCCPDS data. In addition, transactions that are reported very late are associated with illogical gain/loss patterns and other RCCPDS problems.

The date when a transaction is added to RCCPDS is known as its tapedate. The age of a newly reported transaction is measured as the difference between the transaction's tapedate and its effective date.

There is a positive relationship between the existence of an illogical gain/loss pattern and the length of time it takes for a transaction to appear on RCCPDS. This relationship is not, however, strong enough to imply a decision rule that would by itself effectively clean the data.

Although an illogical gain/loss pattern is far more likely to include very late transactions, not all illogical transactions are late and not all late transactions are illogical. Overall, 4.0 percent of transactions are older than six months, and 1.3 percent older than one year. Individuals whose first transaction was part of an illogical pattern had a first or second transaction older than six months for 27.8 percent of the time, versus 4.9 percent for all other individuals with two or more transactions. Similarly, those with three or more transactions on the same date had transactions over six months old 34.8 percent of the time, versus 5.6 percent for all other individuals having three or more transactions.

Most old transactions belong to logical gain/loss patterns. There is nothing apparently wrong with 94.0 percent of the FY 1981 through FY 1985 individual patterns, but 0.7 percent of these include a transaction older than one year. The validity of these records might also be suspect simply due to their age. If some unknown programming or data-recording error can generate an erroneous transaction that produces an illogical pattern, the same mechanism could generate apparently logical but erroneous transactions.

CONCLUSION

The RCCPDS transaction file is a valuable source of data on Navy enlisted personnel in the Selected Reserve. It provides historic files on individual gains, losses, and reenlistments to SELRES that are not available on other data sets.

The greatest weakness of the RCCPDS transaction file is the generation of erroneous transactions. There is some type of transaction inconsistency for 6.0 percent of individuals on the FY 1981 through FY 1985 files. For 2.1 percent of individuals, there is an illogical pattern to their SELRES gain and loss dates. This should not seriously bias most estimates of duration of SELRES service. The one exception may be the group of individuals who leave SELRES and later return. Erroneous transactions may place too many individuals in this group.

Another weakness of RCCPDS is the absence of some data elements applicable only to Navy SELRES. Where possible, the identification of SELRES affiliation program and NEC should be done with the assistance of other data sources.

RCCPDS has several advantages over the IEMF for determining SELRES affiliation and loss dates:

- Availability. RCCPDS transaction files are available from the beginning of FY 1981. Files for previous years also exist, but are of lower data quality.
- Convenient Format. Individual SELRES affiliation and loss transactions have already been created from the IEMF. To create a similar file from scratch would require matching each monthly IEMF (around 200,000 records) to the preceding month's IEMF.
- Fewer Problems with Late Reporting. RCCPDS avoids many of the problems with late processing of SELRES gains and losses by determining an effective date from IEMF data elements rather than from the date the transaction first appeared. Data reporting lags still limit the use of the most recent data.

On balance, the RCCPDS transaction file, despite its problems, is a useful source of data. When possible, Navy SELRES files should be used to supplement the RCCPDS data element, but RCCPDS is a good source of information on individual gain and loss dates in the Selected Reserve.

APPENDIX A

OPTIONS FOR RCCPDS DATA CLEANUP

APPENDIX A

OPTIONS FOR RCCPDS DATA CLEANUP

There is no completely objective way to prepare the RCCPDS transaction data. Because the mechanism through which erroneous transactions occur is not understood, these transactions cannot be systematically identified and eliminated.

Some set of decision rules is necessary if RCCPDS transactions are used to measure SELRES survival, even if the rule is to drop the suspected data. Different sets of rules may be necessary for studies with different objectives. Any set of decision rules should incorporate some combination of the following procedures:

1. Eliminate records for individuals with inconsistent patterns.
2. Impose a logical sequence on transaction patterns.
3. Eliminate late reported transactions.

Rule (1) is a very conservative approach to the problem. It would bias later estimates only to the extent that some group may be disproportionately eliminated.

Using rule (2) is relatively safe when the only problem is transactions that are exact duplicates of each other. It is more of a cause for concern when the effective dates of the transactions are different. The fact that a second SELRES gain is generated after the first does not ensure that it is the second that was erroneously generated.

Elimination of transactions reported a long period of time after their effective date may be justified as a way to eliminate erroneous transactions not otherwise detected. After some period of time, it becomes far more likely that a transaction was erroneously generated rather than just clerically delayed. The danger in this approach is that valid transactions could be deleted while leaving other transactions present for the same individual.

In a data sample containing only gain or loss transactions, elimination of all transactions older than one year reduced the number of individuals in the sample by 0.7 percent and total transactions by 1.7 percent. The proportion of illogical patterns falls by only 5.0 percent, but simultaneous transactions fall by 24.3 percent of their original proportion.

APPENDIX B
RCCPDS DATA FORMATS

APPENDIX B

RCCPDS DATA FORMATS

The following table contains data formats for enlisted personnel in the RCCPDS Transaction Files for FY 1979 through FY 1985. FY 1986 and FY 1987 files each have different formats to accommodate an expansion of RCCPDS data elements.

TABLE B-1

RCCPDS DATA FORMATS: NAVY ENLISTED PERSONNEL

<u>Data element</u>	<u>Location</u>	<u>Length</u>
Reserve component	1	1
Social Security number	2	9
Name	11	18
Reserve component category	29	2
Pay grade	31	2
Date of pay grade	33	4
Full time support status	37	1
Active guard/reserve indicator	38	1
Date of birth	39	6
End of term of service (ETS)	45	6
State of residence	51	2
Zip code	53	5
Pay entry base date	58	6
Rating	64	7
Duty rating	71	7
Source	78	1
Race	79	1
Ethnic group	80	1
Sex	81	1
Marital status	82	1
Number of dependents	83	1
Civilian education	84	1
Professional military education	85	1
Language	86	2
Total active federal military service	88	3
Years of service for retirement	91	2
Program element code	93	6
Unit identity code	99	7
Unit state	106	2
Unit zip code	108	5
Basic branch	113	2
Current aeronautical rating	115	1
Current flying status	116	1
Eligible for 20 year retirement	117	1
Incentive status date	118	4
Incentive status code	122	2
Date of initial military service	124	4
Date of initial reserve service	128	4
Civilian occupation code	134	4
Mental group	138	1
Obligor status	139	1
Term of enlistment	140	1
Disputed data indicator	141	1
Transaction code	142	2
Separation program designator	144	3
Character of service	147	1
Transaction effective date	148	6

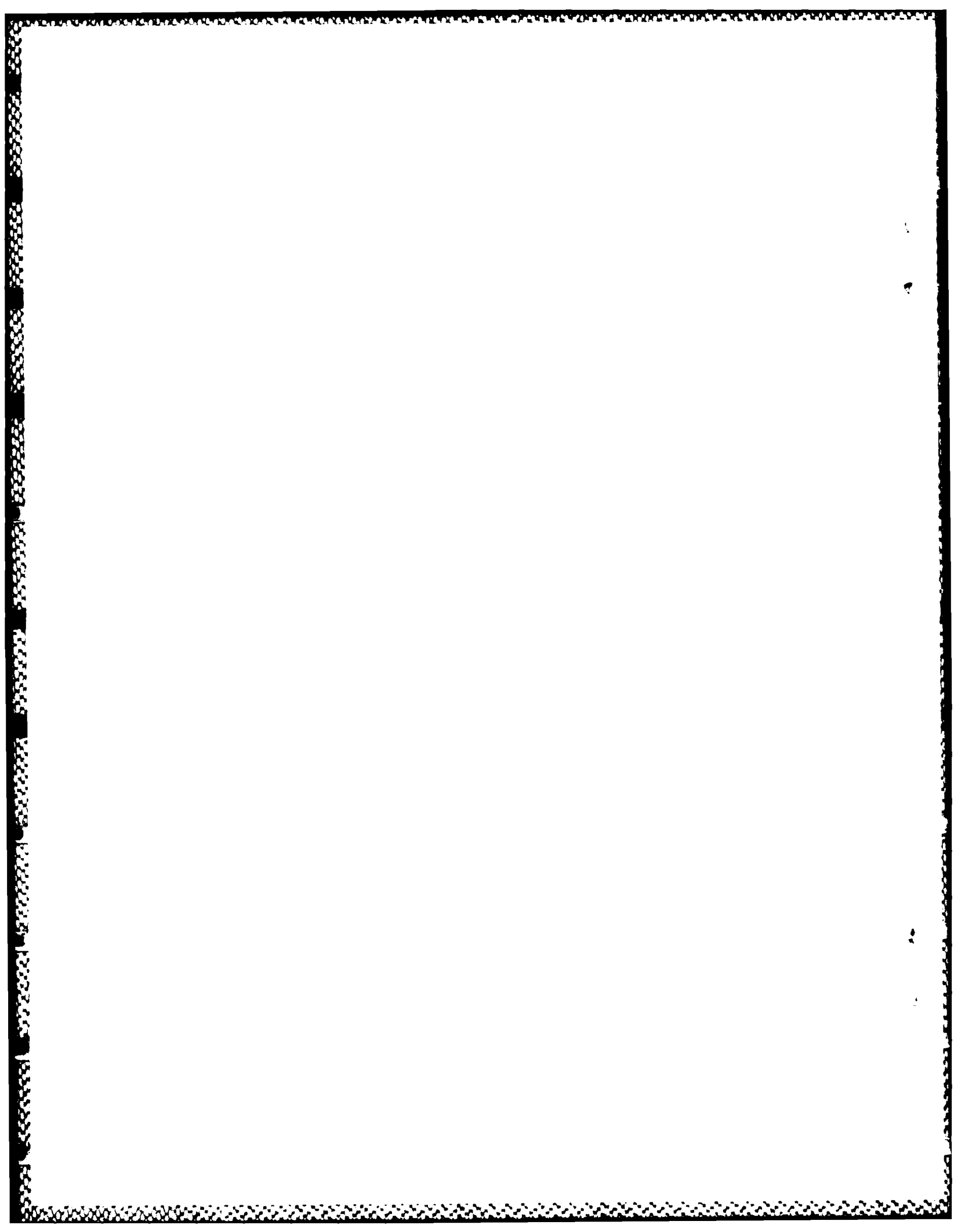
APPENDIX C
NOTES ON RCCPDS DATA ELEMENTS

APPENDIX C

NOTES ON RCCPDS DATA ELEMENTS

The following notes contain miscellaneous information concerning RCCPDS data elements that might be of use to future researchers. The information was gained through use of RCCPDS and conversations with RCCPDS database managers:

- Reserve Component Category. With the exception of new coding for Retired Reservists, the first position of this data element is identical to the Reserve Forces Category (RFC) on the IEMF that identifies SELRES, Individual Ready Reserve, and other Reserve categories. The second position is identical to TCAT for SELRES members. When the RFC is missing on the IEMF, DMDC assigns the code for the first position based on TCAT.
- Source. This is the source of an individual's first enlistment in any military service. For those with one or no term of military service, this is useful in attempts to identify SELRES affiliation program. However, since FY 1984 it has often been miscoded for non-prior-service recruits. About 40 percent of NPS recruits have Source codes of '2,' indicating an active-duty veteran, and transaction codes of 'G3,' indicating a prior-service SELRES gain from civilian life.
- Ethnic Group. Ninety-eight percent of non-blank entries are coded as "other," "none," or "unknown."
- Unit Identification Code. Although this is usually present, it is sometimes coded with JCS UICs rather than Navy UICs.
- Transaction Effective Date. This field is seldom missing on RCCPDS tapes after FY 1979. However, 1.3 percent of transaction effective dates are more than one year older than the date when the transaction was received by RCCPDS. Another 0.8 percent of effective dates appear more than a year after the date they were reported.
- Tapedate. This variable is coded in binary and does not explicitly appear on RCCPDS formats. Tapedate is the year and month when a transaction was added to RCCPDS. DMDC currently places it where the language data element should be.



END

DATE

FILMED

DTIC

JULY 88