NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

LENGTH-OF-SERVICE / SURVIVAL PROFILES METHODOLOGY FOR THE ROYAL AUSTRALIAN NAVY

by

Daniel Dodds

December 2018

Thesis Advisor:	Marigee Bacolod
Second Reader:	Samuel Crawford (Royal Australian Navy)

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** December 2018	**3. REPORT TYPE AND DATES COVERED** Master's thesis	
4. TITLE AND SUBTITLE LENGTH-OF-SERVICE / SURVIVAL PROFILES METHODOLOGY FOR THE ROYAL AUSTRALIAN NAVY			**5. FUNDING NUMBERS**
6. AUTHOR(S) Daniel Dodds			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Using transactional data from Royal Australian Navy (RAN) HR systems from 2002 to 2018, this thesis produces length-of-service/survival profiles for the RAN Officers and Sailors with a focus on testing for heterogeneity in survival profiles along their characteristic dimensions of rank, workforce category/workgroup, gender, age group and calendar year of enlistment cohort. These survival profiles will inform the RAN about the probability of a member remaining in the service at any point in their period of service.

The dataset contained 21,820 periods of service comprising enlistment and separation events (or enlistment and final population snapshot events) and represented 21,495 RAN individuals.

I determined that all analyzed characteristics were important for predicting RAN separation behavior. In particular, females were significantly more likely to separate than males in the early years (up to 10 years of service). Sailors were more likely to separate than Officers. Across the whole RAN, the first year separation rate is 10.3%, which represents the initial matching period for new recruits.

Two methodologies for preparing survival profiles were compared and validated. Kaplan-Meier was found to be the best for single characteristic variable models, and the Cox proportional hazards model was found to be the best for multiple variable specifications.

14. SUBJECT TERMS survival analysis, length of service, survival profile, survival function, hazard function			**15. NUMBER OF PAGES** 159
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

LENGTH-OF-SERVICE / SURVIVAL PROFILES METHODOLOGY FOR THE ROYAL AUSTRALIAN NAVY

Daniel Dodds
Lieutenant Commander, Royal Australian Navy
BS, University of New South Wales (ADFA), 2000

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 2018

Approved by: Marigee Bacolod
Advisor

Samuel Crawford
Second Reader

Yu-Chu Shen
Academic Associate, Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Using transactional data from Royal Australian Navy (RAN) HR systems from 2002 to 2018, this thesis produces length-of-service/survival profiles for the RAN Officers and Sailors with a focus on testing for heterogeneity in survival profiles along their characteristic dimensions of rank, workforce category/workgroup, gender, age group and calendar year of enlistment cohort. These survival profiles will inform the RAN about the probability of a member remaining in the service at any point in their period of service.

The dataset contained 21,820 periods of service comprising enlistment and separation events (or enlistment and final population snapshot events) and represented 21,495 RAN individuals.

I determined that all analyzed characteristics were important for predicting RAN separation behavior. In particular, females were significantly more likely to separate than males in the early years (up to 10 years of service). Sailors were more likely to separate than Officers. Across the whole RAN, the first year separation rate is 10.3%, which represents the initial matching period for new recruits.

Two methodologies for preparing survival profiles were compared and validated. Kaplan-Meier was found to be the best for single characteristic variable models, and the Cox proportional hazards model was found to be the best for multiple variable specifications.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

ACOL	Annualized cost of leaving
ADF	Australian Defence Force
ADFA	Australian Defence Force Academy
AFQT	Armed Forces Qualification Test
CN	Chief of Navy
CNA	Center for Naval Analyses
DE	Direct entry
DFR	Defence Force Recruiting
DNWR	Directorate of Navy Workforce Requirements
GE	General Entry
GFC	Global Financial Crisis
IMPS	Initial Minimum Period of Service
LOS	Length of Service
ML	Maritime Logistics
MLO	Maritime Logistics Officer
NWRF	Navy Workforce Requirements Forecast
OLS	Ordinary Least Squares
RAN	Royal Australian Navy
STEM	Science, Technology, Engineering and Math
UE	University Entry
WG	Workgroup
YOS	Years of Service

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

A. SEPARATION RATES (PROBABILITY OF SEPARATION)

This thesis develops a methodology for preparing survival and hazard function profiles for the Royal Australian Navy (RAN). The survival function relates to the probability of remaining (also known as retention) in the service, while the hazard function relates to the probability of separation from the service. The profiles inform the RAN of a group or individual's propensity to separate from the service at any point in their careers and can take into consideration any of their characteristics including gender, rank, age group, cohort year of entry and workgroup, or any combination thereof.

Figure A is a survival function profile (developed using a Cox Proportional Hazards Model) for the whole Navy. The curve represents the probability of remaining in the service (on the vertical axis) across the cumulative length of service (on the horizontal axis).

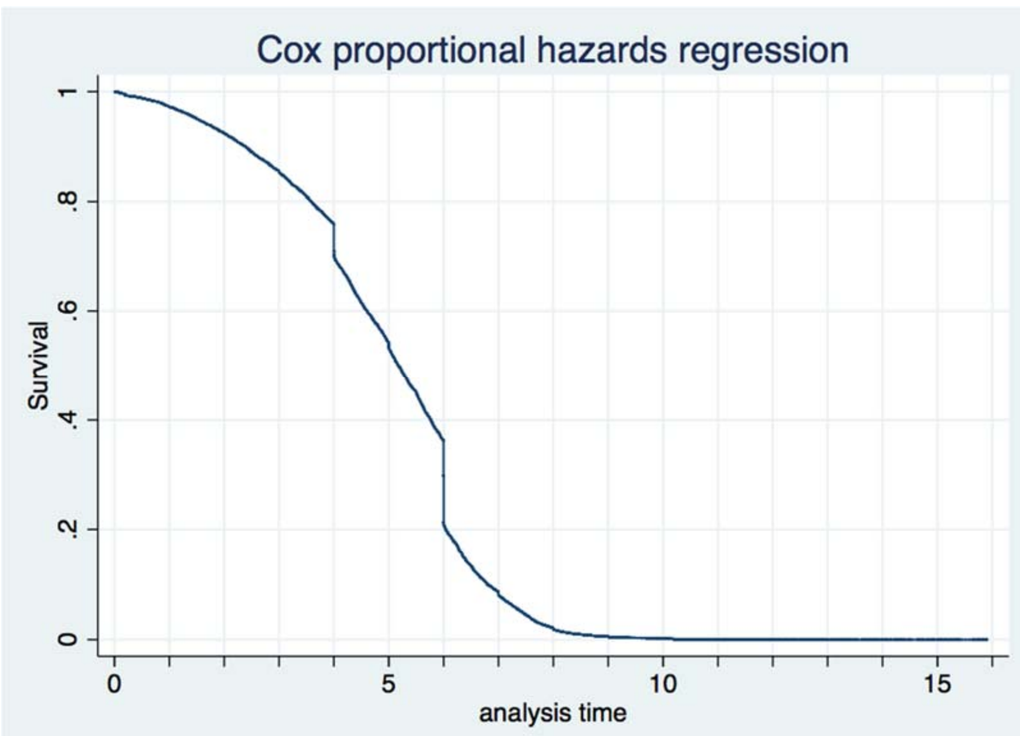


Figure A. Cox survival function across the whole Navy.

The separation rates vary across cumulative length of service (in years). I found that across the whole RAN, the highest likelihood of separation occurs among those who have not yet completed one year of service, with a separation rate of 10.3%, followed by six years of service (6.6%) and then four years of service (4.9%). These latter time periods align perfectly with the RAN Initial Minimum Period of Service (IMPS) "contract" periods, while the early separation indicates potentially bad matches for new RAN members just learning about life in the Navy.

However, I also found that Officers and Sailors have very different separation rates over time. The Sailors' highest three separation rates in order were at zero (11.2%), six (7.2%), and four (5.2%) years of service. The Officers' highest three separation rates in order were at zero (5.8%), one (5.4%) and two (4.1%) years of service. I also found that Officers are more likely to serve for longer periods than Sailors. Figure B is a survival function (developed using Kaplan-Meier model) which shows that in addition to different separation rates across a career between Officers (in blue) and Sailors (in red), the periods of service with the highest instantaneous drops do not match between the two groups. This confirms that Officers service differs significantly from Sailors. This can be partially explained not only by differences in contract periods, but also by the fact that Officers serve for more years on average than Sailors.

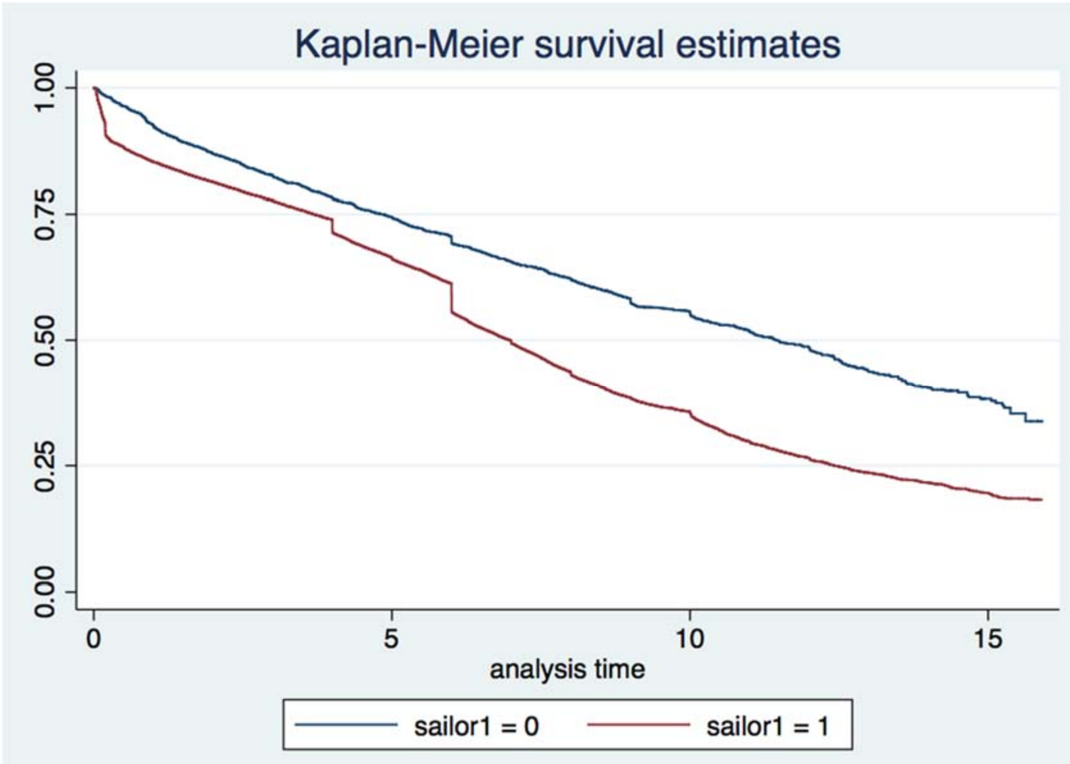


Figure B. Kaplan-Meier survival function (by Officer and Sailor).

B. CHARACTERISTICS OF SEPARATION

All of the analyzed characteristics (i.e., gender, age group, rank, cohort year of enlistment, and workgroup) were investigated as predictors of separation, and all were found to significantly explain separation behavior of RAN personnel.

The aggregated workgroups (RevisedFunction) at enlistment and separation were both found to account for separation behavior of RAN personnel. The traditional lower level workgroup (e.g., "MLO" for Maritime Logistics Officers) was too granular and needed to be consolidated to include all related workgroup entries, in other words, the RevisedFunction of "Maritime Logistics Officers" consolidated the Functions of "MLO," "MLO-T" and "MLO-UT." This added power and significance to the analysis because the sample size in each aggregated workgroup was then increased.

I determined that rank does account for separation behavior of RAN personnel, especially Sailor rank at enlistment and separation, in addition to Officer rank at separation.

Both gender and age group were found to have non-linear relationships with length of service, meaning that the relationship effect varies over a career. Females are 10% more likely on average to separate than males across the whole analysis time. However, and more importantly in the early years of service (2–9 years of service), females are (14–82%) more likely to separate than males, whereas in later years of service females are statistically no different from males with the same years of service. That gender matters more in the earlier years could also coincide with gender differences across the life cycle (e.g., early years are prime childbearing and child rearing years). Figure C displays the survival curve (using the Cox model) by gender and confirms that men (represented in red) and women (in blue) serve differently and this affects their probability of separation.

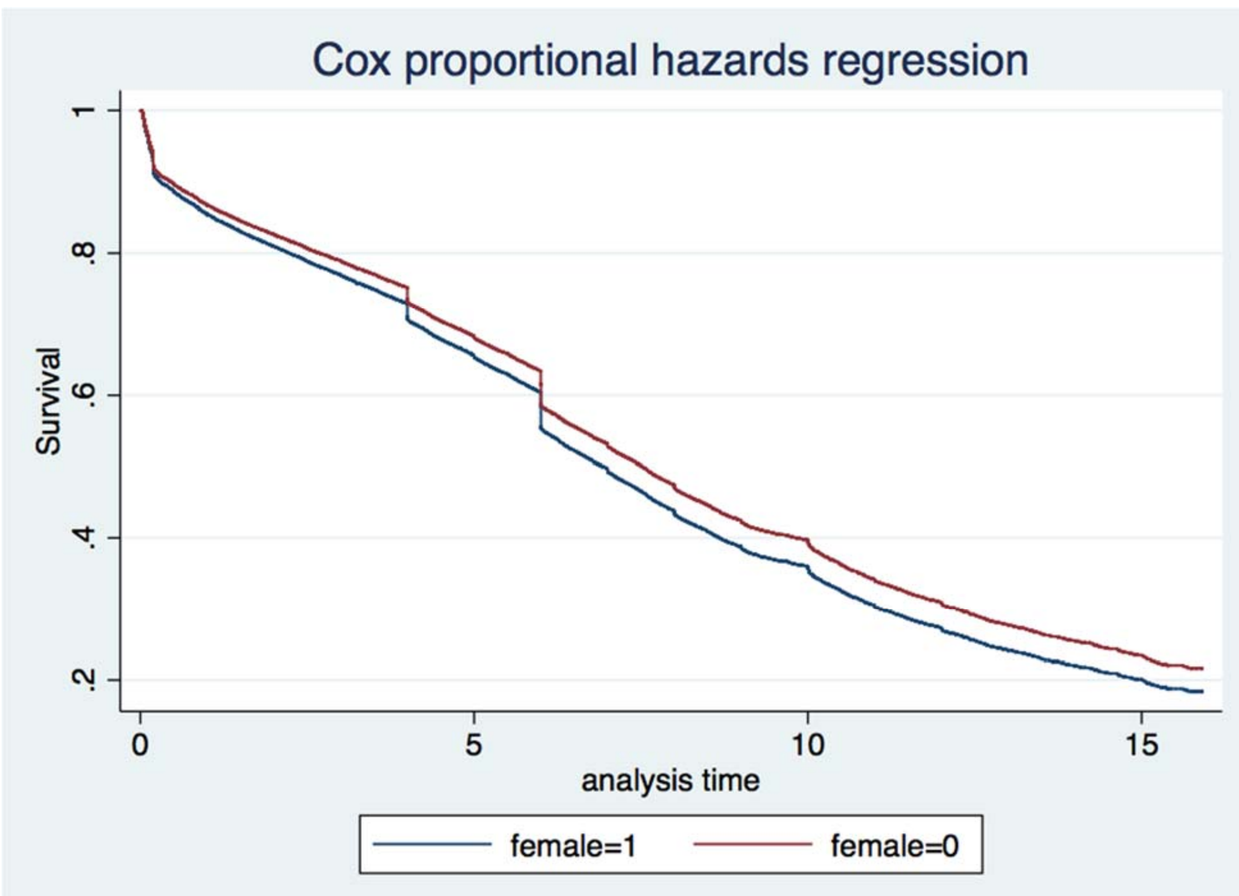


Figure C. Cox survival function by gender.

Figure D shows a gender density analysis (the proportion of women in each workgroup against the proportion of their workgroup in the total Navy population) that was conducted. This shows that Science, Technology, Engineering and Math (STEM) fields have very low female density compared to total proportion of the workgroup in the Navy. As an example, the "Marine Technician" workgroup contains over 15% of the Navy personnel. However, it is made up of less than 1% females.

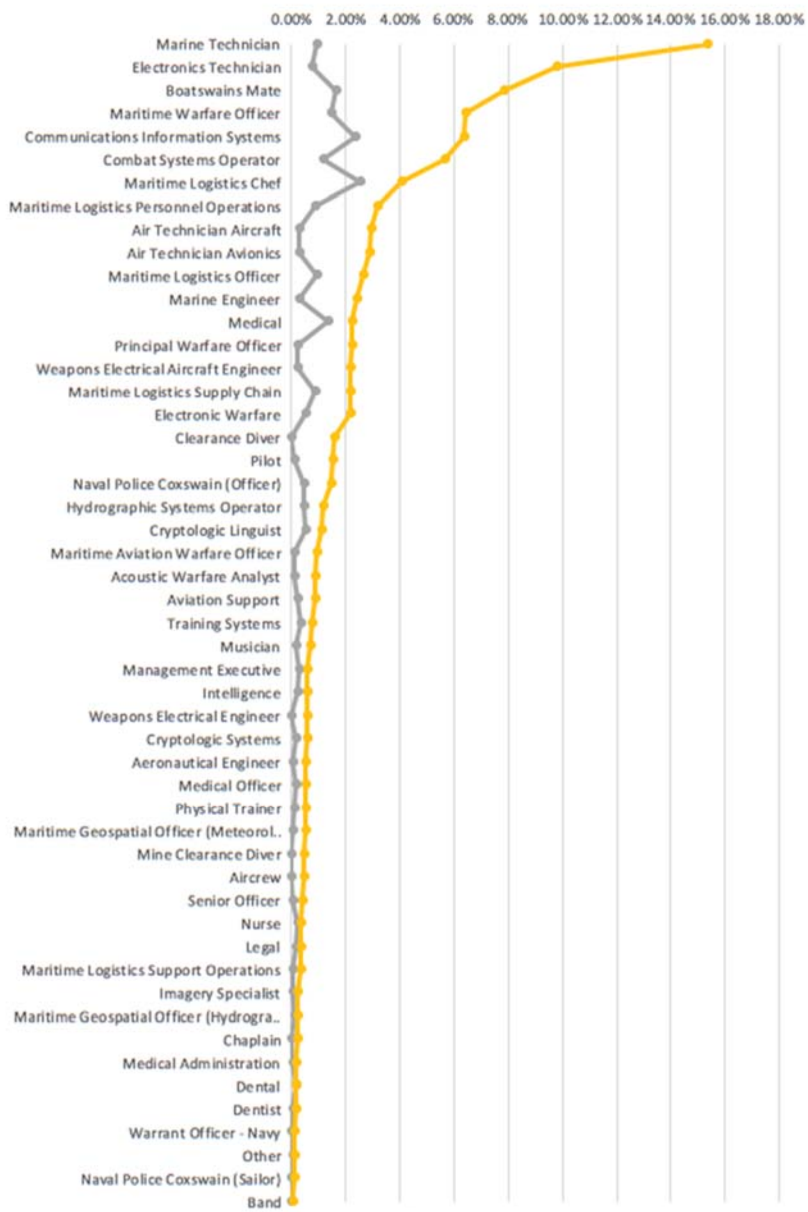


Figure D. Gender density analysis by RAN workgroup.

Analysis of the calendar year of enlistment was conducted to determine whether cohort and/or peer effects are relevant to the separation behavior of RAN personnel. Two effects were found to be at play: the cohort effect and an economic conditions effect. The cohort effect saw that separation is less likely for later cohorts as each new year begins. The economic conditions effect found that in comparison to 2008 (the year of the Global Financial Crisis (GFC) in Australia), the cohort years 2002–2007 indicate that personnel are more likely to separate than the baseline year, and for the cohort years 2009–2018 indicate personnel are less likely to separate. Therefore, the cohort year of enlistment may be a good proxy for macroeconomic conditions.

I found that the youngest age group (16–19 years old) were such outliers in separation behavior compared to all other age groups primarily because they were the largest enlistment age group and had the highest separation rate.

C. FUTURE RESEARCH RECOMMENDATIONS

For further research, I would recommend detailed analysis of the relationship between RevisedFunction0 (and to a lesser extent RevisedFunction1) interacted with other key variables in this analysis (gender, rank, age, and cohort year of enlistment). This would allow the analyst to see, for example, whether the gender differences in separation behavior in early years hold only for certain workgroups or across RAN as a whole. The significant gender differences in separation behavior in the early years of service contrasted with no difference in later years is particularly interesting. Further study could flesh out a more complete picture and aid in formulating gender-specific manpower policies.

I also recommend a further investigation of Officer versus Sailor survival curves using a non-proportional methodology, for example, Kaplan-Meier models or a Machine learning Random Forest algorithm methodology. The Random Forest algorithm is more computationally demanding; however, it may be able to get at these nonproportional differences more accurately and is worthy of investigation.

I also recommend a further investigation be undertaken into other pre-enlistment individual quality characteristics as an avenue to resolve the identified high first year separation rate in addition to the youngest age group effects on separation behavior. Doing this is likely to enable the RAN to determine whether the early turnover rate is acceptable and whether there are other pre-enlistment characteristics which may better predict early separation.

Finally, I recommend that an investigation is conducted to determine whether using the outputs of a Cox model provides significantly better forecast results than another form of predicting separation rates. Utilizing the outputs from a survival analysis and bringing that into a separation simulation model like a Markov may pay large dividends than the current models used.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I am forever indebted to my thesis advisors, Dr. Marigee Bacolod and Lieutenant Commander Samuel Crawford, RAN, for their wise counsel and guidance during the development of this thesis. A huge thank-you to you both for all of your unwavering assistance and support.

I would like to thank the Royal Australian Navy, my sponsor, for this amazing opportunity.

I would also like to thank the Naval Postgraduate School and in particular the faculty and staff of the Graduate School of Business and Public Policy. I have learned a tremendous amount from you all. In addition, I appreciate all of the tireless work undertaken by the International Graduate Programs Office.

Finally, I would like to thank and acknowledge my husband, James, for his dedication, patience, persistence and support during my time at NPS.

THIS PAGE INTENTIONALLY LEFT BLANK

I. ROYAL AUSTRALIAN NAVY BACKGROUND

A. ROYAL AUSTRALIAN NAVY IN THE CONTEXT OF WORKFORCE PLANNING

The Royal Australian Navy (RAN) is the maritime arm of the Australian Defence Force (ADF), commanded by the Chief of Navy (CN). The RAN "provides maritime forces that contribute to the Australian Defence Force's capacity to defend Australia, contribute to regional security, support global interests, shape the strategic environment and protect national interests" (Royal Australian Navy, n.d.-a, sec. "About the Royal Australian Navy").

The RAN has continually suffered significant workforce shortfalls for several decades, limiting its effectiveness in operations and increasing the stress on the existing workforce. In 2014 the Australian National Audit Office (ANAO) conducted an audit on the "Recruitment and Retention of Specialist Skills for the RAN" (Australian National Audit Office, 2014). During this audit the ANAO found that "long-standing personnel shortfalls in a number of 'critical' employment categories had persisted, and the RAN had largely relied on retention bonuses as a short to medium term retention strategy" (ANAO, 2014, para. 11). These 'critical' employment categories are yet to be resolved, indicating deficiencies in the traditional manpower planning practices of the RAN.

This thesis will provide a previously unexplored method for analyzing the RAN workforce, allowing greater insight into the behavior of its constituents, to better inform policy decisions enabling more effective planning and retention of the RAN workforce.

1. Mission, Vision and Motto

The mission of the RAN is "to fight and win at sea" (RAN, n.d., sec. "Mission"), the vision is "An Australian Navy renowned for excellence in service to the nation" (RAN, n.d., sec. "Vision") and the motto is "Serving Australia with Pride" (RAN, n.d.-a, sec. "Vision").

2. Roles and Functions

To achieve these requirements, the RAN can conduct “maritime patrol and response, interdiction and strategic strike, protection of shipping and off-shore territories and resources, maritime intelligence collection and evaluation, and escort duties” (RAN, n.d., sec. “About the Royal Australian Navy”) in addition to peacetime activities, which “include maritime surveillance and response within Australia’s offshore maritime zones, hydrographic, oceanographic and meteorological support operations, humanitarian and disaster relief, and maritime search and rescue” (RAN, n.d.-a, sec. “About the Royal Australian Navy”).

3. Organizational Structures

The RAN is structured into two primary commands, Strategic Command and Fleet Command. The former is responsible for the strategic direction of the RAN (RAN, n.d.-b) while the latter conducts all of the raise, train, and sustain of Navy ships and operational units, effectively managing the Fleet-in-being (RAN, n.d.-c, para. 1.10). The RAN currently operates 49 commissioned ships (including amphibious assault ships, destroyers, frigates, landing ship dock, mine hunters, oiler replenishment vessels, patrol boats, submarines, hydrographic and survey vessels and other support ships), three non-commissioned ships, as well as helicopter squadrons, clearance diving teams, other assorted commands and units. In addition to the current force, the RAN is expanding substantially with the roll out of new destroyers, frigates, submarines, offshore patrol vessels, and supply support (tanker) ships.

The RAN is an all-volunteer force and is manned by approximately 14,000 men and women.

4. Ranks and Rates

Officers join as Midshipman and are promoted from Acting Sub Lieutenant (O1) to Admiral (O10). Enlisted Sailors join the RAN as a Recruit and once qualified are promoted from Seaman (E1) to Warrant Officer of the Navy (E10). Figure 1 shows the Officer and Sailor ranks.

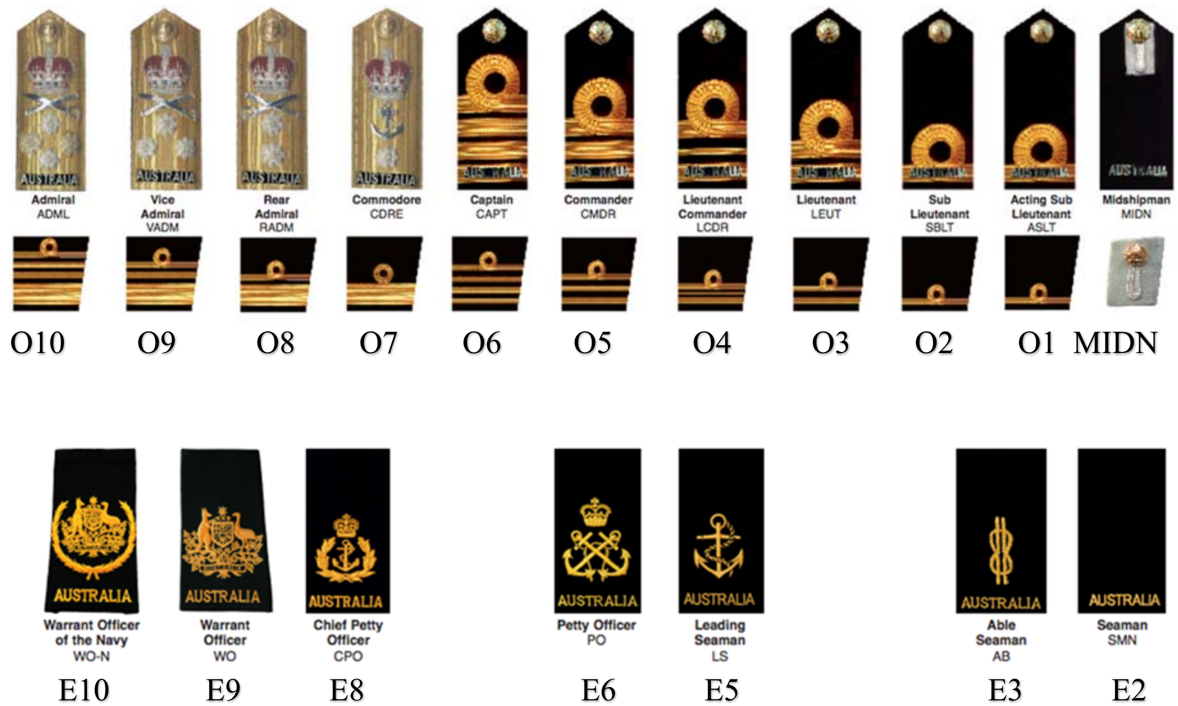


Figure 1. Royal Australian Navy Officer and Sailor ranks. Adapted from Royal Australian Navy (n.d.-d).

5. Workgroup Structures

RAN personnel are grouped into communities or families based on their functional employment categorization. These higher-level groupings are aviation, chaplain, engineering, health services, logistics & administration, senior officer and warfare. Each grouping is further separated by family. For example, part of the logistics & administration grouping is the Maritime Logistics (ML) family and this would contain all ML related Officer and Sailor categories including Maritime Logistics Officers (MLO), Maritime Logistics - Supply Chain (ML-SC), Personnel Operations (ML-P), Chefs (ML-C) and Support Operations (ML-S) Sailor categories. And finally, each category of Officer and Sailor is called a Workgroup (WG); for example all of the Maritime Logistics Officers would be grouped together into one WG. There are 15 Officer WGs called Primary Qualifications (PQ) and 26 Sailor WGs called categories (or rates). Table 1 shows an example breakdown of the community structures.

Table 1. Royal Australian Navy community structures

Grouping	Family	Workgroup
Logistics & Admin	Maritime Logistics (ML)	Maritime Logistics Officer (MLO)
		Maritime Logistics - Supply Chain (ML-SC)
		Maritime Logistics - Personnel Operations (ML-P)
		Maritime Logistics - Chef (ML-C)
		Maritime Logistics - Support Operations (ML-S)
Aviation	…	…

The central document relating to this is the Australian Navy Publication (ANP) 2110 - RAN Career Management (this is a RAN internal publication not available to the public). Each WG (and most families) has a range of policy documents and guidelines that details how each is to be planned, managed and executed. These include information pertaining to career management, sea/shore ratio, career pathways, promotions eligibility, time in rank requirements, key career continuum positions and training requirements. (RAN, n.d.-c).

Meanwhile, different classes of RAN ship are manned with a range of WG personnel with differing proportions of each depending on the mission, roles and functions. The sea/shore ratio of personnel is the ratio of the time spent at sea on a ship versus the time spent working in a shore establishment (Navy land base). It is usually measured in years, and the exact ratio varies across (and within) WGs.

6. Workforce (Manpower) Requirements

The Directorate of Navy Workforce Requirements (DNWR) is the section within the Navy responsible for Navy workforce requirements planning and analysis. DNWR produces the Navy workforce requirements forecast (NWRF) in collaboration with the WG managers and acquisition projects. The NWRF gathers the estimated manpower requirements of each acquisition project, and compares that demand against the expected breakdown of each WG to develop a breakdown of the manpower requirements, forming the authorized baseline against which forecasting and modelling is performed. DNWR is also responsible for analyzing the sustainability and affordability of current and future workforce requirements. DNWR uses the NWRF to model the recruiting and promotion targets necessary to meet the forecasted manpower requirements. The results of this body of work are aggregated into the RAN recruiting directive (RAN, n.d.-c).

7. Recruiting

Defence Force Recruiting (DFR), a joint command, is responsible for delivering the targeted number of enlistees to the ADF through attraction, screening and processing of applicants to all three services (Navy, Army and Air Force). DFR is provided with annual recruiting directives from each service with targets for the recruiting requirements for that year. DFR works with its uniformed, public service civilian and prime contractor staff to achieve these targets. Prospective recruits are processed through a series of tests to determine their eligibility for each service and WG. Differing standards apply for each WG according to the skills applicable to their respective roles. For example, the Electronics Technician WG involves roles that requires maintenance of highly complex electrical equipment, necessitating a higher standard of academic, aptitude, physical and psychometric requirements than that of WGs that do not include complex technical roles, such as the Boatswain's Mate WG.

8. Entry Methods

General Entry (GE). This is the method of entry predominantly employed for any Sailor joining the RAN. Applications for a Sailor WG who have no prior experience in military service and have enhanced skills, such as previous diesel fitting trade or electrician certificates, will enter via this method.

Direct entry (DE). Applicants who apply for a commission to become an Officer in the RAN have the option to apply either directly to a WG career, or to spend their initial career studying for a degree at the Australian Defence Force Academy (ADFA). Undergraduate education is considered highly valuable to Officer applicants. Therefore, if a desirable applicant does not possess a sufficient level of undergraduate education, but meets all other requirements to be an Officer, they may be encouraged to proceed through the ADFA entry before rejoining Direct Entry applicants in their respective WG career pipelines.

B. CHAPTER SUMMARY

This thesis will develop a novel methodology for examining the recruitment, promotion and retention policies of each WG with respect to each constituent's probability of remaining in relation to their length of service by employing survival analysis, a method that has yet to be deployed to the RAN manpower planning process. This analysis will assist in determining the effects of various RAN manpower policies on the RAN workforce to better inform manpower policy decisions.

II. ACADEMIC LITERATURE REVIEW

In this chapter I provide a review of the academic and scholarly military literature relevant to this study. First, Section A provides a broad overview of defense manpower economics. Then, Section B discusses specific details regarding defense manpower planning. Finally, Section C outlines the econometric methods used in the defense manpower analysis field.

This chapter will conclude with detail on how this research fits into the current span of knowledge and demonstrates how these frameworks apply to the RAN context.

A. ECONOMICS OF MANPOWER

Asch, Hosek, and Warner (2007) state the four major events which shaped the defense manpower agenda, calling for research "to help policy makers deal with the challenges that these factors presented" (Asch et al., 2007, p. 1076). These events include the end of the Cold War, the rise in U.S. college attendance, the increased operational tempo of U.S. forces, and the rising cost of U.S. military entitlements (Asch et al., 2007).

The supply-demand framework is a primary tool used in economics to describe the rational behavior of its actors. I will then describe these four major events through the prism of the supply-demand framework for manpower.

1. Demand for Labor

As stated in Asch et al. (2007) the main event impacting the demand for manpower in most Western militaries was the end of the Cold War. The official cessation of the Cold War resulted in a significant military downsizing across many NATO, European and Western countries including the U.S. (The World Bank, n.d.). The U.S. alone underwent a 38 percent reduction in military manpower strength (Asch et al., 2007). In addition, the reduction in demand (and therefore military force manpower requirements) led many countries to eliminate conscription and switch to all-volunteer forces (Asch et al., 2007). Implementing an all-volunteer force model was a significant change for many countries not just in how they attracted and recruited new personnel but in how they managed them when in the service.

2. Supply of Labor

The supply of labor available to U.S. armed forces was impacted by a number of trends concurrent with the end of the Cold War. As reported by Asch et al. (2007) the first trend was the increased attraction of a college degree compared to military service which led to an increase in college attendance, thereby increasing the difficulty of recruiting (Asch et al., 2007).

The decision to join the military or to reenlist is explained via the standard occupational choice theory from Rosen (1986) and as adapted by Warner and Asch (1995). The pay and non-pecuniary benefits in each sector (both military and civilian) are compared to arrive at an individual's decision (Warner and Asch, 1995). This theory is based on comparing the utility (maximization) function for each sector (U^M for military and U^C for civilian), considering the military wage (W^M) and civilian wage (W^C) and the non-pecuniary benefits (τ^M and τ^C) (Warner and Asch, 1995). People choose to join the military only when the following equalities are true:

$$U^M = W^M + \tau^M > U^C = W^C + \tau^C \qquad (or) \qquad W^M - W^C > \tau^C - \tau^M.$$

Some of the non-pecuniary benefits include the value of serving one's country, "risk of injury or death" (Asch et al., 2007, p. 1078) along with the individual's "tastes for military and civilian service" (Asch et al., 2007, p. 1078).

The second factor which Asch et al. (2007) describe is related to the long-lasting improvement in the U.S. economy which led to the "civilian unemployment rate fall to its lowest level in 30 years … from 7% in 1992 to 4% in 2000" (Asch et al., 2007, p. 1078). This saw a further tightening in the supply of youth due to the growing economy and rising employment which created increased competition for the target youth (Asch et al., 2007).

The third factor relates to the "population of veterans who positively recommend military service" (Asch et al., 2007, p. 1079). This argument highlights the fact that the U.S. population of veterans has been declining since the Cold War and therefore if that population of good word-of-mouth advertising for the U.S. military also declines then this would result in reduced supply of potential recruits (Asch et al., 2007).

Additional trends resulting from the Cold war manning and resource downsizing were changes in available policy levers including military pay, retirement, bonuses and allowances and also ones specifically relating to recruiting--advertising, recruiters, educational benefits, and other recruiting resources (Asch et al., 2007). The reduction in military funding effectively saw declines in all of the incentives to join the military that directly resulted in a reduction in the attractiveness (and utility) of military service and therefore supply (Asch et al., 2007).

Finally, the increased operational tempo, "combat operations and fatalities in Afghanistan and Iraq had negative … effects on high-quality enlistments" (Asch et al., 2007, p. 1085). The increased public visibility of combat fatalities saw more individuals swing towards civilian employment over military service which has taken a sizable toll on Army recruiting (Asch et al., 2007).

All of these supply of labor trends impacted the way the U.S. (and other Western countries) attracted, recruited and retained personnel (Asch et al., 2007). It is likely the RAN was also impacted in a similar fashion by these trends, due to the mirroring nature of Australian society and its own similar relationship to the Cold War.

In terms of individual characteristics, a probit or logit model can be used to determine the characteristics associated with enlistment decisions (Asch et al., 2007). Using a probit model, Kilburn and Klerman (1999) found that for high school seniors, a "negative effect of AFQT (Armed Forces Qualification Test) score, mother's education, family income, and a positive effect of number of siblings" on the propensity to enlist (Asch et al., 2007, p. 1081). For high school graduates they found "a negative effect … of civilian wages, being unemployed … and having children" while "marriage had a positive influence" (Asch et al., 2007, p. 1081).

In addition to the two choice model, Kilburn and Klerman (1999) used a multinomial logit model with a third choice, the decision to attend college (Asch et al., 2007). They found a negative effect from AFQT score, age when a senior, and mother's education and a positive effect from "variables associated with availability of resources to pay for college" (Asch et al., 2007, p. 1082).

3. Recruitment

In addition to studies that broadly fit within the supply-demand framework for manpower, there are studies focused explicitly on recruiting policies. The U.S. services target "the recruitment of high-quality youth, defined as high school diploma graduates who score in the upper half of the AFQT distribution" (Asch et al., 2007, p. 1078). Various studies have found that advertising, recruiter productivity, and recruiter incentives are critical to achieve the recruiting mission (Asch et al., 2007). A number of studies reviewed in Asch et al. (2007) conclude that a range of enlistment bonuses and incentives have been very effective in attracting more high-quality recruits as well as in recruiting individuals into more difficult functional employment categories.

4. Retention and Attrition

The development of retention theory was "spurred by the need to account systematically for present and future compensation in the empirical analysis of retention" (Asch et al., 2007, p. 1091). In recent years, retention theory has been incorporated into the "analysis of alternative compensation structures and their role in inducing individuals to remain in the service" (Asch et al., 2007, p. 1091).

There are a range of models which have been used to assist with modeling retention behavior, each with its own set of modeling assumptions, advantages and disadvantages (Asch et al., 2007).

a. ACOL

The Annualized cost of leaving (ACOL) random utility model developed by Warner and Asch (1995) is a "one-period model in which reenlistment depends on a measure of military/civilian pay" (Asch et al., 2007, p. 1091). In this model, the "individual weighs current pay and the stream of future pay for each … pathway under consideration" (Asch et al., 2007, p. 1091). Many variables representing explanatory characteristics can be added to the model to determine the effect of each, including reenlistment bonus, unemployment rate, race, ethnicity, gender and marital status. This model is well suited to cross-sectional data (Asch et al., 2007).

b. ACOL-2

The ACOL-2 model is an extension of the ACOL model in that it introduces separate terms for taste and for transitory shock and it can also be used on panel data (Asch et al., 2007). Taste refers to the preference for the military and transitory shock is when individuals are buffeted in period after period by some event/s (Asch et al., 2007).

c. Dynamic Programming

The first application of dynamic programming to retention was the work by Gotz and McCall (1984) which developed a "theoretical model of Air Force officer retention and to empirical estimates of its parameters" (Asch et al., 2007, p. 1092). This model incorporates the value at a certain time of having entered the military at a prior time period, along with the value at a certain time of having stayed in a civilian job that started at a prior time period, and finally the individuals' personal discount factor (Asch et al., 2007).

The model can also take into consideration the value of leaving, the probability of promotion, demotion, rank and year of service in addition to any policy rules (Asch et al., 2007).

Extensions to this model by Asch and Warner (2001) saw the addition of individual ability and individual effort, formulating the output as "expected years of service" and "retention of high ability individuals" (Asch et al., 2007, p. 1094). For empirical tractability, however, there are a number of modeling assumptions; in particular, the sources of heterogeneity are reduced (Asch et al., 2007).

As stated in Asch et al. (2007) the "probability of a particular choice, given one's present state, has a multinomial logit form" (p. 1093) and using the dynamic programming retention model, the probability of staying in the military for one period (or multiple) can be determined.

5. Force Management (i.e., Management of Current Force in Being)

"Manpower is a key ingredient in the production of military readiness" (Asch et al., 2007, p. 1105). Without effective manpower resources a country's military would not be able to achieve its missions and this would result in a dire impact to the country's ability to defend itself as well as negatively impacting the economy.

The goal of force management is "to continually balance the force along the multiple dimensions of experience, quality, skills and active/reserve/civilian status to achieve the most readiness at least cost" (Asch et al., 2007, p. 1105). Achieving the most efficient force possible is the goal of any manpower planner. To get the best value for money is critical.

According to Asch et al. (2007) and Warner and Asch (1995) "experience is a key component of productivity" and "quality matters" (Asch et al., 2007, p. 1106). As an individual's experience increases so does effectively their personal productivity; in addition, their unit also becomes more productive because of the mentoring and training benefits. The studies find that higher performance is also linked to higher education and entrance scores, hence quality does matter (Asch et al., 2007).

Therefore, to benefit from a more productive and higher performing group one must ensure that there is a good balance of experience (measured in years of service, and skills) and quality (higher education and entrance scores) (Asch et al., 2007).

The trends and changes stemming from the end of the Cold War resulted in amendments to the existing military manpower models to attempt to cover all of these environmental factors discussed above (Asch et al., 2007).

B. MANPOWER PLANNING, LENGTH-OF-SERVICE PROFILES

1. Closed Labor Market

A "'closed labor market' is the foundation of military personnel management" (Rodney, 2017, p. 27). It is closed in that recruitment is the only avenue for entry, promotees are grown and selected from within the organization and there are usually minimal options for lateral entry from other foreign militaries, other services, or in-service transfers from another category (Rodney, 2017). The majority of military organizations are completely closed labor markets; however, some militaries are attempting to deal with declining recruitment pools by attempting to attract other skilled and experienced applicants via mid-level entry from other militaries or from the private sector.

The RAN has predominantly employed a closed labor model. However, recently the RAN introduced the Total Workforce Model (TWM), which seeks to enable personnel inflow into RAN service at desirable later points in their career, as a means to supplement hollow workforce areas and provide greater flexibility to workforce managers.

The primary challenge for military workforce (manpower) planners operating in a closed labor market is to determine how many personnel are required--including the breakdown by rank, category, experience, skills, etc.—to ensure that assigned missions can be achieved and the organization can be sustainable (Rodney, 2017).

Generally people join the military with zero experience or relevant skills and gain those skills and experience over time, and according to Rodney (2017) years of service (YOS) is used to measure experience (Rodney, 2017). However, Rodney fails to address many variables that are highly likely to impact experience, such as a measure of speed of promotion, number of early promotions, level of education, types of training, performance in high-profile or high-stress positions, etc. Many of these would be widely available for analysis, and their inclusion in the study is necessary to minimize potential bias in any results.

2. Length-of-Service Distribution

Rodney (2017) produced the following example of a Length of Service (LOS) distribution by number of personnel using the following equation.

$$P \text{ with } n \text{ YOS in year } t \leq P \text{ with } n\text{-}1 \text{ YOS in year } t\text{-}1;$$

$$P \text{ represents the number of personnel.}$$

Figure 2 displays this LOS distribution and this will identify the total number of personnel at each YOS at a snapshot in time (Rodney, 2017).

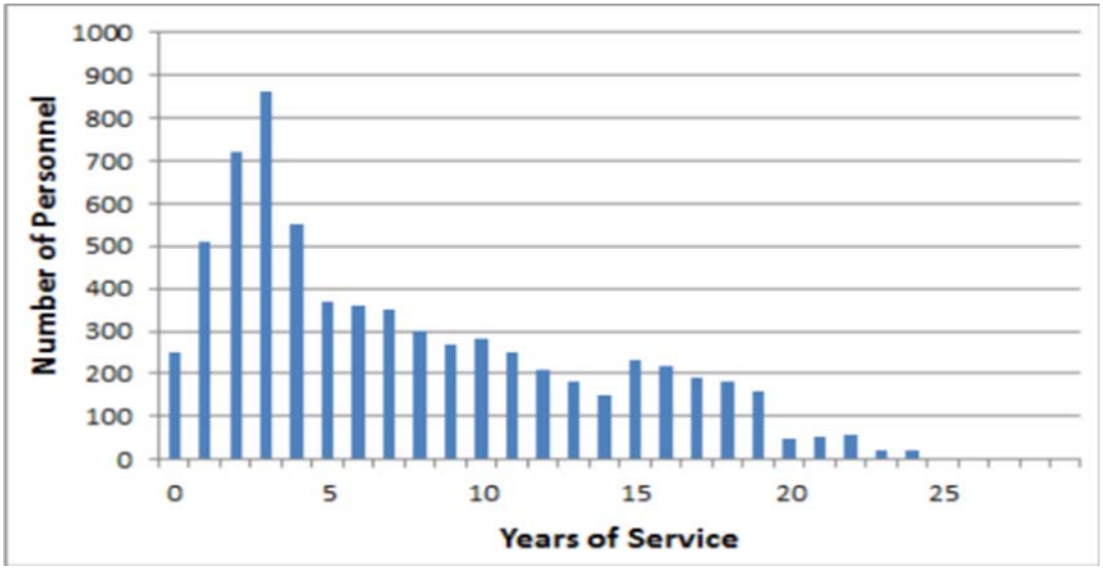


Figure 2. Length-of-service distribution for U.S. Navy operations specialists rating. Source: Rodney (2017).

The LOS distribution provides a great insight into the level of experience in an organization (Rodney, 2017). It does tell us something about the studied group in that we can see the effects of changes in recruiting requirements (mission) (Rodney, 2017). Larger recruiting requirements increases the number of personnel in a certain YOS group, but at the same time decreases can result in a smaller YOS cohort/s (Rodney, 2017). This distribution can also be used as a simple snapshot of attrition (Rodney, 2017).

To achieve a manning state of 50 personnel with five YOS you need to have at least 50 personnel with four YOS (Rodney, 2017). To forecast how many personnel will be within each YOS bucket you need to calculate the following:

$$\#\ of\ personnel\ with\ 5\ YOS = Continuation\ rate * (\#\ of\ personnel\ with\ 4\ YOS).$$

Rodney (2017) explains that the continuation rate is the "number of personnel who remain from one year to the next" (Rodney, 2017, p. 29) and that this is calculated as the percentage proportion of personnel with four YOS expected to transition to five YOS (Rodney, 2017). These calculations can be iterated to determine the number of personnel who remain or 'survive' for more than just a single year snapshot (Rodney, 2017).

3. Survival Curves (Profiles)

Workforce (manpower) planners are regularly asked to determine quantity and quality (education, experience, skills) of personnel in future periods (Rodney, 2017). For example, if the Navy needs 10 personnel with a particular skill each year for the next 10 years, how does the Navy achieve this and maintain it?

A survival curve is a suitable option to answer questions of this nature (Rodney, 2017). These curves are closely related to LOS distributions covered above and they are constructed with the percentage of accessions from year one which remain on active duty (survive) against each YOS (Rodney, 2017).

To prepare the survival curve, we must know the personnel strength (inventory) at a point in time (t) and with n YOS (Rodney, 2017). To calculate the future inventory, we need to multiply the current known inventory by the proportion of inventory that we know will continue to serve into the following period and increase their YOS by one (Rodney, 2017). This proportion is the continuation rate $CR(n)$ (Rodney, 2017). To state this mathematically (Rodney, 2017, p. 29):

$$INV(n+1, t+1) = INV(n,t) * CR(n).$$

This formula can be iterated over multiple periods to achieve the following (Rodney, 2017, p. 29):

$$INV(n+2, t+2) = INV(n,t) * CR(n) * CR(n+1), \ldots \text{etc..}$$

So, to calculate the personnel strength at any point in the future we multiple the original recruited figure (X) by the continuation rate for each period from enlistment up until the desired future date. This is stated as (Rodney, 2017, p. 30):

$$\text{Future strength at } n \text{ years} = X * CR(0) * CR(1) * \ldots CR(n-1),$$

$$\text{where } 0 <= n <= \text{max service period.}$$

The survival curve depicted in Figure 3 is the proportion of the future strength of the original recruited figure (as a percentage) plotted at each YOS (Rodney, 2017).

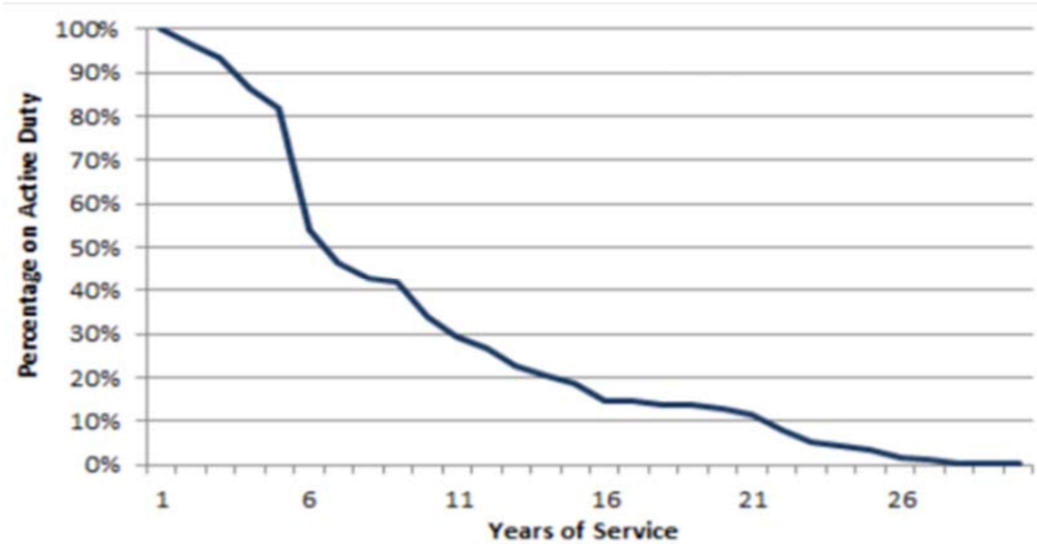


Figure 3. Survival curve. Source: Rodney (2017).

These survival curves are not conditional on anything. Conditionality means that the function needs to meet the mandatory requirements (conditions) before its valid. For example, before I make bread I need flour and water, but if I don't have water I can't make bread. But if I have all of the ingredients (the conditions) the bread can be made and is an output of this process.

C. METHODS

1. Logistic Regression Model (Multinomial Logit Form/Logistic Function)

Logistic regression is similar to ordinary regression in that here is "a dependent variable, *y*, and one or more independent variables" (Anderson, Sweeney, Williams, Camm, & Cochran, 2018, p. 725). Unlike other regressions the dependent variable is coded as a binary variable (meaning 0 or 1) and the coefficients on the equation independent variables relates to the probability of the dependent variable occurring (Anderson et al., 2018). Unlike ordinary regression, a logit is also a nonlinear regression.

For example, in Asch et al. (2007), a logit or probit model of individuals' decisions to enlist is used, with demographic (age, family background) and environmental characteristics (location) as predictors for the probability eligible youth would select the military versus civilian opportunities. They are also “useful for targeting recruiting effort” (Asch et al., 2007, p. 1081).

Logistic regression “can be applied to estimate the probability of experiencing a particular lifetime event within a limited time period; nevertheless, it does not consider the time when the event occurs and therefore disregards the length of the survival process” (Liu, 2012, p. 3).

Logistic regression uses a logit link function:

$$f(\mu_Y) = ln(\frac{P}{1-P}).$$

2. Probit Regression Model

Probit regression is very similar to logistic (logit) regression, with the main difference being that it uses an inverse normal link function (Liu, 2012):

$$f(\mu_Y) = \Phi^{-1}(P).$$

The output is interpreted in the same way that a logit model is.

3. Survival Analysis

According to Cleves, Gould, and Gutierrez (2004), survival analysis “concerns analyzing the time to the occurrence of an event” (p. 1). Liu (2012) describes survival analysis as the “use of reason to describe, measure, and analyze features of events for making predictions about not only survival but also ‘time-to-event processes’ … such as from living to dead, from single to married, or from health to sick” (p. 1). It is known by many names, depending on the field its being applied in and is ‘time to event analysis’ and ‘duration analysis’ in economics, ‘reliability theory/analysis’ in engineering, and ‘event history analysis’ in sociology. I will use the term survival analysis in this thesis for consistency purposes.

An event in the survival analysis context is defined as the event/s which are of concern to the analysis (Cleves et al., 2004). In the original medical context, examples of these events include the death of a cancer patient or onset of a disease or remission, but in economics the event could be the moment an individual gets a job (becomes employed) when they have been unemployed for a period of time (Cleves et al., 2004).

The time or duration is the time period which has passed until the event occurs from a certain starting time or period (Cleves et al., 2004).

Covariates are additional variables which are added to the model to attempt to explain the variation in time to an event (Cleves et al., 2004).

A survival function, $S(t)$, is the probability of surviving past time *t*; In other words, "it is the probability that there is no failure event prior to *t*" (Cleves et al., 2004, p. 7). It equals 1 at time $t=0$ and reduces as t goes to infinity. It is also the reverse cumulative distribution function of T (Cleves et al., 2004):

$$S(t) = 1 - F(t) = Pr(T > t).$$

The hazard function, $h(t)$, is the "probability that the failure event occurs in a given interval, conditional upon the subject having survived to the beginning of that interval" (Cleves et al., 2004, p. 7). The function is an "instantaneous rate of failure" (Cleves et al., 2004, p. 7) and is described as:

$$h(t) = \lim_{\Delta t \to 0} \frac{Pr(t+\Delta t > T > t | T > t)}{\Delta t} = \frac{f(t)}{S(t)}, or$$

$$h(t|x) = h_0(t) * \exp(x\beta).$$

The hazard function can take on values from zero (no risk) to infinity (failure will occur at that instant in time) and the rate can also change over time (Cleves et al., 2004).

A hazard ratio is an odds ratio and can be interpreted when the ratio is greater than one something is more likely to happen, vice a ratio less than one means something is less likely to occur. The hazard ratio has the following equation:

$$HR = \frac{h(t|x)}{h(t|x=0)} = e^{(x\beta)}, where\ x\ represents\ the\ covariate.$$

One can calculate all the other functions (e.g., probability density function, cumulative distribution function, and survivor function) from having at least one of them (e.g., hazard function) through algebraic manipulation (Cleves et al., 2004).

“In real data we often do not observe subjects long enough for all of them to fail” (Cleves et al., 2004, p. 2). An example of this is if an individual signs up to the start of a medical drug trial, receives multiple drug treatments over a period of 6 months, but then doesn’t complete the full course of treatment for an unknown (non-death related) reason. This is classified as censoring. The researchers did not observe this individual for the entire trial period, but there is still valuable data for this individual that can be analyzed. Survival analysis allows us to model datasets without having to remove these incomplete observations (Cleves et al., 2004).

Why not just use simple ordinary least-squares (OLS) linear regression, logit or probit? One of the key assumptions of OLS linear regression is that the residuals (errors) are normally distributed (Wooldridge, 2016). Regularly survival data will not meet this requirement and therefore another method was required; “At its core, survival analysis concerns nothing more than making a substitution for the normality assumption characterized by OLS with something more appropriate for the problem at hand” (Cleves et al., 2004, p. 2). Using a more realistic residuals distribution assumption leads to parametric survival analysis (Cleves et al., 2004).

Previous analysis methods (logit/probit) required complete longitudinal data and this would reduce the effectiveness of these methods on many data sets without complete observations (Cleves et al., 2004). It also led to analysis not being conducted on the most recent of data, because the full life cycle was not observed during the period.

Survival analysis is an extension of logistic regression, which is iterated at each time interval to achieve a conditional analysis.

a. Kaplan-Meier (1958) (Non-parametric)

“When no covariates exist, or when the covariates are qualitative in nature (gender, for instance), we can use nonparametric methods such as Kaplan and Meier … to estimate the probability of survival past a certain time or to compare the survival experiences for each gender” (Cleves et al., 2004, p. 5).

Kaplan and Meier (1958) produced the Kaplan-Meier estimate which is "a nonparametric estimate of the survivor function" (Cleves et al., 2004, p. 93) and it also accounts for censoring and other survival analysis features (Cleves et al., 2004). Kaplan-Meier also makes no assumptions about the "distribution of the failure times nor how covariates serve to shift … the survival experience" (Cleves et al., 2004, p. 5).

b. Cox Proportional Hazard Model (1972) (Semi-parametric)

Unlike Kaplan-Meier, the model developed by Cox (1972) does allow consideration of covariates (Cleves et al., 2004). The Cox model also "produce estimates … such as probability of survival past a certain time" (Cleves et al., 2004, p. 6) and it also requires "a conditional logistic model is fit for each analysis" (Cleves et al., 2004, p. 4).

Liu (2012) states that the "major contribution of the Cox model, given its capability of generating simplified estimating procedures in analyzing survival data, is the provision of a flexible statistical approach to model the complicated survival processes as associated with measurable covariates" (Liu, 2012, p. 3).

The fact that "results of analyses are being determined by the assumptions and not the data is always a source of concern" (Cleves et al., 2004, p. 3). At the time this problem resulted in a search for alternate techniques which did not involve distribution of failure time assumptions, meaning that the model will allow the data to select the best assumption (Cleves et al., 2004).

Finally, for the Cox model to be able to analyze multiple covariates a proportional assumption was required. Effectively, this meant that any variable/characteristic had to be a multiplicative factor of its baseline hazard or survival function, meaning that they were parallel to each other.

D. CHAPTER SUMMARY

In this chapter I have summarized the current state of literature for the economics of Defense manpower, manpower planning and finally the econometric methodologies utilized in the Defense manpower area of research. Neither logit and probit models are able to handle conditional analyses; and Kaplan-Meier can only analyze single covariates, therefore survival analysis is proposed as the best available option for the analysis of RAN workforce survival and this will be confirmed and validated in Chapter IV.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DATA AND METHODOLOGY

A. INTRODUCTION

This chapter describes the dataset in detail along with the variables contained therein. In addition to this the descriptive statistics for the full RAN population snapshot (as of 31 May 2018) is developed to place the future findings and results into the context of the actual RAN workforce characteristics. Finally, a summary of the methodology to be used in this thesis is provided.

B. DATA

The sponsoring organization, DNWR, has provided the dataset for this thesis. The dataset contains pre-collected movement transaction data from RAN Human Resources (HR) systems from the period 1 Jul 2002 to 31 May 2018. The data has been stripped of Personally Identifying Information (PII) by DNWR prior to this analysis. Each observation in the dataset represents an enlistment or separation event for an individual Permanent Navy (active duty) RAN Officer or Sailor along with their characteristics at the time of the event. The transactional level variables (i.e., event details) contain event date, event type, event reason, event comment, method of enlistment and/or separation type. The individual's characteristics (variables) include rank, gender, age group (i.e., 20–24 years old) and workgroup/employment categorization variables (grouping, revised function, family, function, sub function, and skill grade).

The full dataset (prior to dropping irrelevant events) includes 182,271 observations.

Conducting this study will provide the RAN insight into its current probabilities of retention at any point in a member's career. Based on the outcomes of the survival analysis, and length of service/survival profiles, the RAN will be able to pinpoint critical points in a service members' career based on all of the dimensions which may impact retention and enable policy change to correct these issues.

C. VARIABLES

Following is an enumeration of the variables in the original data set and from which I constructed analysis variables. With each I describe what the variables mean and how I constructed analysis variables.

(1) ID

This variable is a consecutive number assigned to each individual represented in the data set. This number is used in lieu of the individual's service number (PMKEYS number).

(2) Age Group at Event Date

This indicator variable represents the individual's age interval category at the event date. These categories include age groups 16–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64 and 65–69 years of age.

For the purposes of the Survival analysis, this variable was further adapted into two indicator variables representing the age group at enlistment (AG0) and the age group at separation/current status (AG1).

(3) Gender

This variable (Gender) describes the individual's gender as female 'F' or male 'M'.

In addition to this variable, indicator variables were created for female and male.

(4) Event Date

This variable (EventDate) contains the actual event date relevant to each event recorded in the HR system. For example, for an individual's enlistment event, the event date would be the date of enlistment into the RAN.

For the purposes of the duration analysis methodologies this variable was further separated into event date at enlistment (date0) and event date at separation/current status (date1).

(5) Event Type

This variable (EventType) contains the type of event relating to the data entries. This includes standardized event categorizations for 'Enlistment', 'Separation' and 'Current Status'. The 'Current Status' event represents a full RAN population snapshot at a specific point in time.

For the purposes of the duration analysis methodologies, this variable was split into event type at enlistment (EventType0) and event type at separation/current status (EventType1).

(6) Event Reason

This variable (EventReason) further clarifies the event type with a standardized reason. For 'Enlistment' these identify whether the individual was recruited as an ab initio (meaning a recruit with no prior RAN experience), overseas recruitment, reenlistment, service transfer from the Australian Regular Army (ARA) or the Royal Australian Air Force (RAAF), transfer from the Gap Year program, or transferred from the reserves. While for 'Separation' this variable describes the primary reason for leaving the service, including unsuitable for service, contract completion, resignation, training failure, medically unfit, etc.

For the purposes of the duration analysis methodologies, this variable was split into event reason at enlistment (EventReason0) and event reason at separation/current status (EventReason1).

(7) Comment1

This variable (Comment1) provides elaborating information that supplements the EventReason. For example, the EventReason may indicate an Enlistment event, and Comment1 will detail the type of enlistment (ab initio, re-enlistment, transfer from reserves etc.).

(8) MOE

This variable (MOE) represents the individual's method of entry into the RAN. For Sailors it identifies their recruit/General Entry (GE) class number, whilst for Officers it identifies their entry avenue: ADFA, Direct Entry (DE), Undergraduate Entry (UE).

(9) MOE Supp

This variable (MOESupp) only states each Sailor's GE class.

(10) Separation Type

This variable (SeparationType) is only relevant to individuals with a recorded separation event and it specifies whether their separation was voluntary or involuntary.

(11) Grouping

This variable (Grouping) states the individual's higher level functional employment grouping. There are seven of these groupings including 'Aviation', 'Chaplain', 'Engineering', 'Health Services', 'Logistics & Admin', 'Senior Officer' and 'Warfare'.

For the purposes of the duration analysis methodologies, this variable was split into grouping at enlistment (Grouping0) and grouping at separation/current status (Grouping1).

(12) Revised Function

This variable (RevisedFunction) represents one level lower in the functional employment groupings. For each Grouping there are numerous aggregated functions. This employment function is basically a job title for each Officer and Sailor's functional employment category (e.g. Maritime Logistics Officer).

For the purposes of the duration analysis methodologies, this variable was split into aggregated revised function at enlistment (RevisedFunction0) and aggregated revised function at separation/current status (RevisedFunction1).

(13) Family

This variable (Family) details the individual's functional employment family. Underneath the grouping level are these 25 short 2–5 digit codes which represent the family. An example is 'ML' which includes all Maritime Logistics Officers and Sailors.

For the purposes of the duration analysis methodologies, this variable was split into family at enlistment (Family0) and family at separation/current status (Family1).

(14) Function

This variable (Function) represents the same information as in the Revised Function, however it states each as a short code which Sailors use as part of their rank (i.e. ABML-SC represents Rank: Able Seaman, Function: ML-SC). For each function it also includes an additional code to represent the under training distinctions.

For empirical tractability, I aggregate Functions to group together those under training with those already in those Functions, as they are practically similar. In the analysis itself there are 69 aggregated Revised Functions and originally 169 Function codes. It is necessary to aggregate these codes given the small sample sizes for some of these detailed Function codes (e.g., there are only n=5 in Function code 'DEN MNGR' representing the Dental Manager Sailor, however in the Revised Function this is now aggregated up to n=51 which includes all other Dental related Sailor categories).

For the purposes of the duration analysis methodologies, this variable was split into function at enlistment (Function0) and function at separation/current status (Function1).

(15) Sub Function

This variable (SubFunction) represents another level of specification for the function. For functional employment categories which have a sub-classification this is where it is recorded.

For the purposes of the duration analysis methodologies, this variable was split into sub-function at enlistment (SubFunction0) and sub-function at separation/current status (SubFunction1).

(16) Skill Grade

This variable (SkillGrade) represents another level of the function. For some functional employment categories the skill grade relates to a lower level sub-classification, whilst for most it represents the pay group specification within an individual's workgroup.

For the purposes of the duration analysis methodologies, this variable was split into skill grade at enlistment (SkillGrade0) and skill grade at separation/current status (SkillGrade1).

(17) Rank Code

This variable (RankCode) contains a three digit code representing each rank level. For Sailors their rank codes are prefixed with 'E' whilst Officer's rank codes are prefixed with 'O'. All rank codes contain two number digits after the letter representing the rank in the hierarchy of ranks. The number starts at '00' and increases with each rank step. For Sailors the rank codes range from E00 to E10, whilst Officers range from O00 to O10.

For the purposes of the duration analysis methodologies, this variable was split into rank at enlistment (RankCode0) and rank at separation/current status (RankCode1). In addition to these the variables were also segregated into Sailor rank code at enlistment (SRank0), Sailor rank at separation/current status (SRank1), Officer rank at enlistment (ORank0) and Officer rank at separation/current status (ORank1).

(18) Additional Generated Variables

Finally, in addition to the above variables a number of additional variables were generated to enable the analyses to be conducted. These include:

1. Numspell—which represents the number of spells (periods of service) for an individual;

2. Censor—an indicator variable which represents whether the spell was censored (i.e., individual was still serving at the end of the data set);

3. Fail—an indicator variable which represents whether the spell failed (i.e., separated prior to the end of the data set);

4. year0—represents the year of the enlistment event;

5. year1—represents the year of the separation/current status event;

6. LOS—which represents the length of service in years for each spell and is calculated by subtracting year0 from year1.

7. cumLOS—which represents the cumulative length of service in years, meaning the total length of service across all spells for an individual. This is calculated by summing all of the LOS per individual;

8. sailor0—which represents whether the individual was a Sailor at enlistment;

9. sailor1—which represents whether the individual was a Sailor at separation/current status;

10. officer0—which represents whether the individual was an Officer at enlistment; and

11. officer1—which represents whether the individual was an Officer at separation/current status.

D. DESCRIPTIVE STATISTICS

1. Gender

I begin by describing the gender composition in the RAN. As of May 31, 2018, the total Navy is comprised of 13,727 personnel. Of this strength, 21% are female and 79% are male, as reported in Table 2 and illustrated in Figure 4.

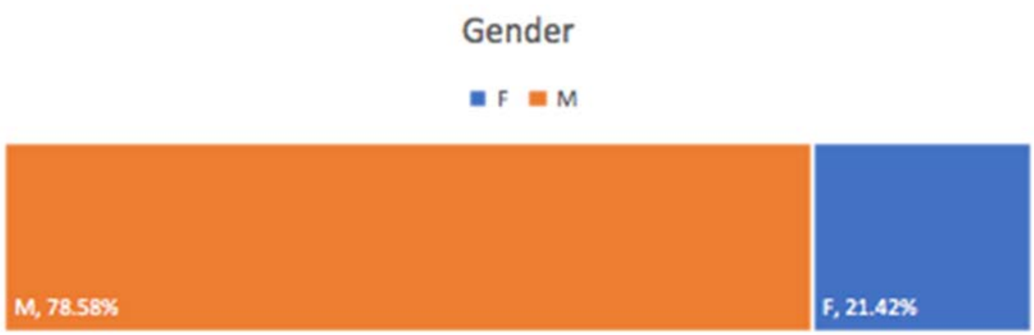


Figure 4. Gender breakdown.

Table 2. Gender breakdown.

Gender	Freq.	Percent
F	2,941	21.42
M	10,786	78.58
Total	13,727	100

2. Rank

In terms of rank, the largest group of RAN Sailors are at the E03 rank code which represents the rank of 'Able Seaman' whist for Officer's that is the O03 rank code which is 'Lieutenant." The breakdown of RAN personnel by rank is reported in Table 3 and displayed in Figure 5.

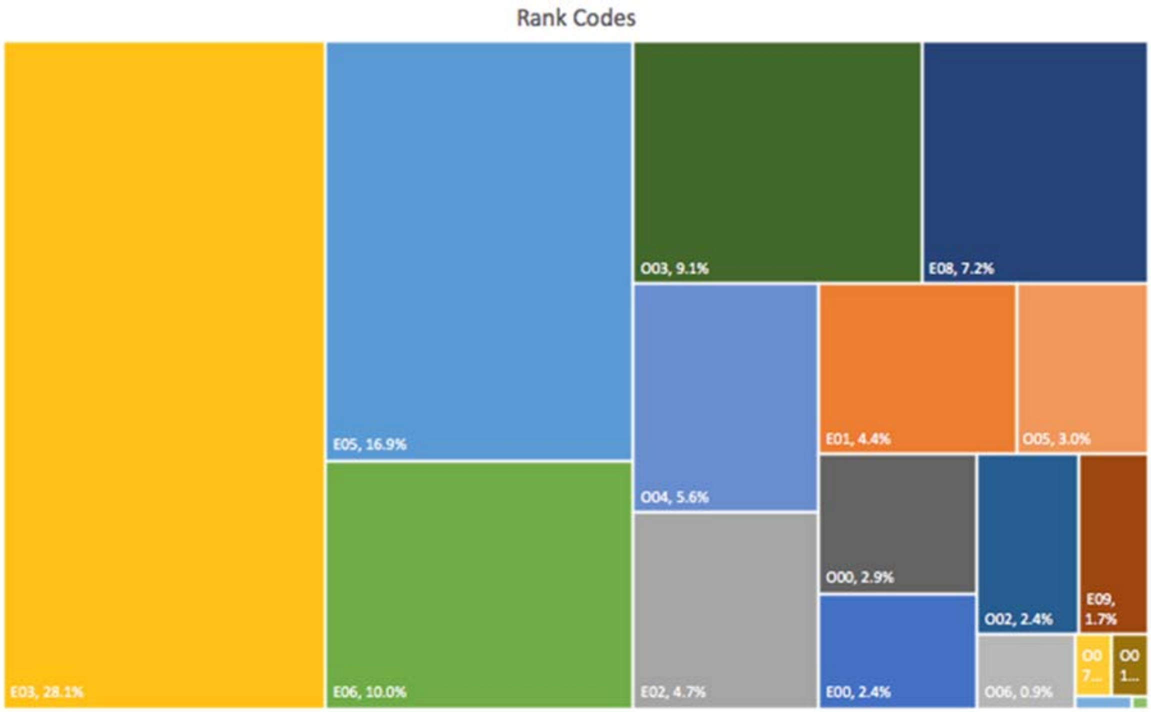


Figure 5. Rank codes breakdown.

Table 3. Rank codes breakdown.

RankCode	Freq.	Percent
E00	328	2.39
E01	607	4.42
E02	650	4.74
E03	3,864	28.15
E05	2,321	16.91
E06	1,371	9.99
E08	984	7.17
E09	228	1.66
E10	1	0.01
O00	399	2.91
O01	40	0.29
O02	324	2.36
O03	1,251	9.11
O04	765	5.57
O05	407	2.96
O06	130	0.95

RankCode	Freq.	Percent
O07	41	0.3
O08	13	0.09
O09	3	0.02
Total	13,727	100

3. Sailors versus Officers

RAN Officers make up 25% of the force and Sailors are 75% (see Table 4).

Table 4. Sailors versus Officers breakdown.

	Freq.	Percent
Sailors	10,354	75.43
Officers	3,373	24.57
Total	13,727	100

4. Technical versus Non-technical Sailors

Breaking down the Sailors even further, technical Sailors represent 41% of the RAN and non-technical Sailors 59% (see Table 5).

Table 5. Technical versus non-technical Sailors breakdown.

	Freq.	Percent
Technical Sailor	4254	41.09%
Non-technical Sailor	6100	58.91%
	10354	100.00%

5. Groupings

The two largest groupings are Warfare and Engineering totaling approximately 70% of the RAN. Table 6 reports the breakdown of RAN personnel by Grouping and is displayed in Figure 6.

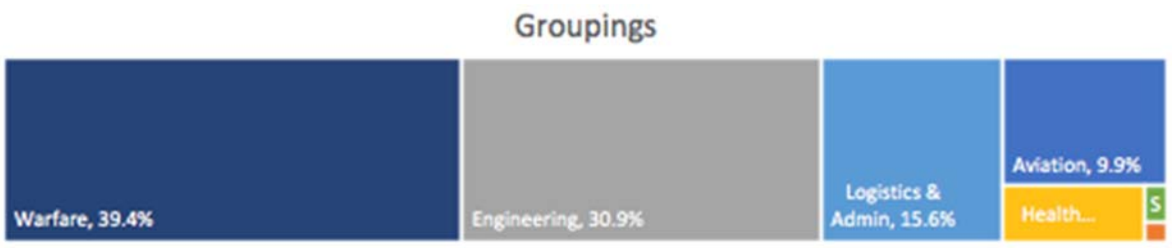


Figure 6. Groupings breakdown

Table 6. Groupings breakdown

	Freq.	Percent
Aviation	1353	9.86%
Chaplain	30	0.22%
Engineering	4,241	30.90%
Health Services	504	3.67%
Logistics & Admin	2,141	15.60%
Senior Officer	55	0.40%
Warfare	5,403	39.36%
Total	13,727	100.00%

6. Workgroup

The largest workgroup is the ‘Marine Technician’, which is a Sailor category responsible for upkeep and maintenance of ship’s mechanical and electrical distribution systems.

The largest Officer workgroup is the ‘Maritime Warfare Officer.’ See Table 7 for the breakdown of RAN personnel by Workgroup.

Table 7. Workgroup (variable: RevisedFunction) breakdown.

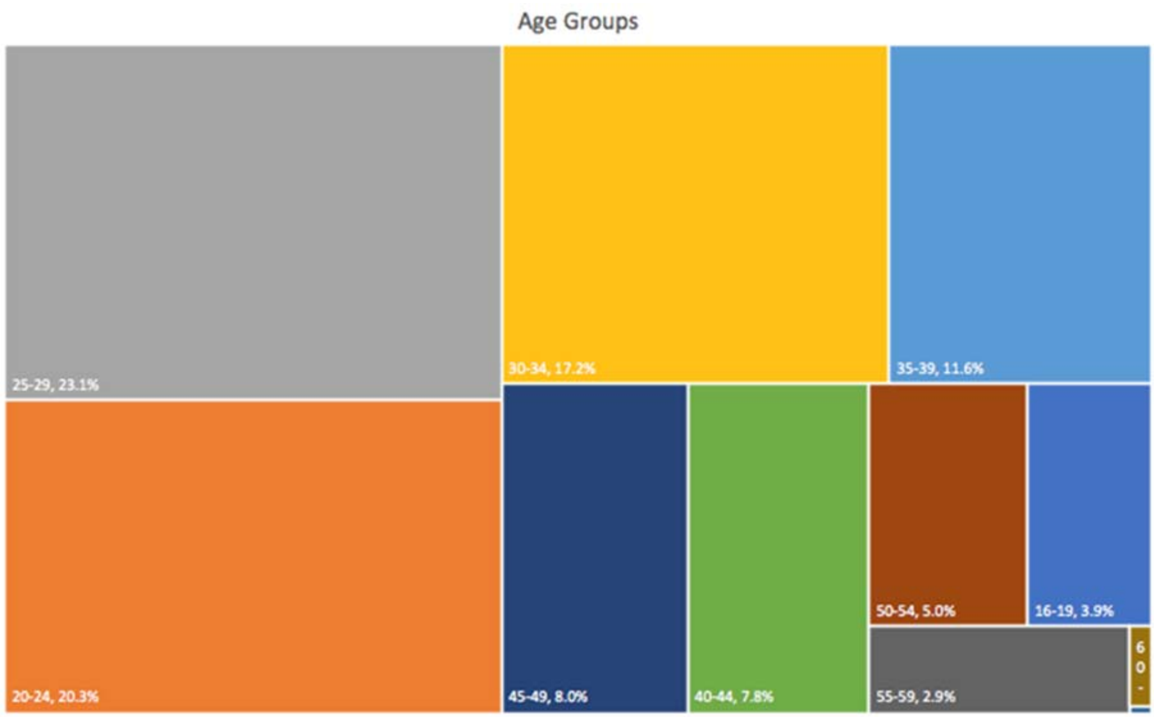
Workgroup	Freq.	Percent
Acoustic Warfare Analyst	119	0.87
Aeronautical Engineer	74	0.54
Air Technician Aircraft	406	2.96
Air Technician Avionics	395	2.88
Aircrew	64	0.47
Aviation Support	118	0.86
Band	6	0.04
Boatswains Mate	1,079	7.86
Chaplain	30	0.22
Clearance Diver	218	1.59
Combat Sys Supervisor A-AIC	2	0.01
Combat Systems Operator	442	3.22
Combat Systems Operator Mine Warfare	107	0.78
Combat Systems Operator-Above	219	1.6
Combat Systems Operator-Above-ASAC	1	0.01
Combat Systems Operator-Under	5	0.04
Communications Information Systems	782	5.7
Communications Information Systems Submariner	89	0.65
Cryptologic Linguist	149	1.09
Cryptologic Systems	77	0.56
Dental	25	0.18
Dentist	24	0.17
Electronic Warfare	200	1.46
Electronic Warfare Submarines	95	0.69
Electronics Technician	1,207	8.79
Electronics Technician Submariner	138	1.01
Hydrographic Systems Operator	158	1.15
Imagery Specialist	34	0.25
Intelligence	83	0.6
Legal	48	0.35
Management Executive	83	0.6
Marine Engineer	295	2.15
Marine Engineer Submariner	33	0.24
Marine Technician	1,881	13.7
Marine Technician Submariner	227	1.65

Workgroup	Freq.	Percent
Maritime Aviation Warfare Officer	126	0.92
Maritime Geospatial Officer (Hydrographer)	72	0.52
Maritime Geospatial Officer (Meteorologist/Oceanographer)	32	0.23
Maritime Logistics Chef	439	3.2
Maritime Logistics Chef Submariner	30	0.22
Maritime Logistics Officer	284	2.07
Maritime Logistics Personnel Operations	259	1.89
Maritime Logistics Supply Chain	360	2.62
Maritime Logistics Supply Chain Submariner	16	0.12
Maritime Logistics Support Operations	302	2.2
Maritime Logistics Support Operations Submariner	12	0.09
Maritime Warfare Officer	788	5.74
Maritime Warfare Officer Submariner	95	0.69
Medical	282	2.05
Medical Administration	27	0.2
Medical Officer	73	0.53
Medical Submariner	24	0.17
Mine Clearance Diver	65	0.47
Musician	100	0.73
Naval Police Coxswain (Officer)	12	0.09
Naval Police Coxswain (Sailor)	205	1.49
Nurse	49	0.36
Other Officers	2	0.01
Other Sailors	15	0.11
Physical Trainer	72	0.52
Pilot	207	1.51
Principal Warfare Officer	302	2.2
Principal Warfare Officer Amphib	1	0.01
Senior Officer	55	0.4
Training Systems	108	0.79
Warrant Officer (Entry)	19	0.14
Warrant Officer - Navy	1	0.01
Weapons Electrical Aircraft Engineer	50	0.36
Weapons Electrical Engineer	299	2.18
Weapons Electrical Engineer Submariner	31	0.23
Totals:	13,727	100

7. Age Groups

The largest age group category is the 25–29 years old. Almost half (47%) of the RAN is aged under 30. Table 8 contains the breakdown of RAN personnel by age group and Figure 7 displays this information.

The RAN allows individuals to join the service at any age provided the applicant can complete their Initial Minimum Period of Service (IMPS) 'contract' prior to the relevant compulsory retirement age.



Note: The brown section in the bottom right hand corner represents age group 60–64 with 0.23% and the smallest section represents age group 65–69 with 0.02%.

Figure 7. Age groups breakdown.

Table 8. Age groups breakdown.

	Freq.	Percent
Age1619	539	3.93%
Age2024	2781	20.26%
Age2529	3165	23.06%
Age3034	2358	17.18%
Age3539	1594	11.61%
Age4044	1070	7.79%
Age4549	1096	7.98%
Age5054	688	5.01%
Age5559	402	2.93%
Age6064	31	0.23%
Age6569	3	0.02%
	13727	100.00%

8. Gender Density by Workgroup

Figure 8 provides an analysis of the gender density by workgroup across the RAN. The total workgroup density (i.e., workgroup proportion of the total Navy—in yellow) is calculated by totaling the number of individuals in each workgroup divided by the total number of individuals in the RAN. The female density for each workgroup (in grey) is calculated in the same way except that the numerator is the total number of women in each workgroup. Showing these two figures side by side illustrate the density of various RAN functions and how the RAN is organized by function, while also showing how gender composition varies across both.

The x-axis in Figure 8 sorts the workgroups by number of people with that the largest of workgroups on the left hand side. The y-axis represents the density percentage.

As demonstrated in Figure 8 the largest RAN workgroups (on the left) have some of the smallest female densities. There are conflicting explanations for low female participation rates in military and STEM fields, and these won't be covered in this thesis.

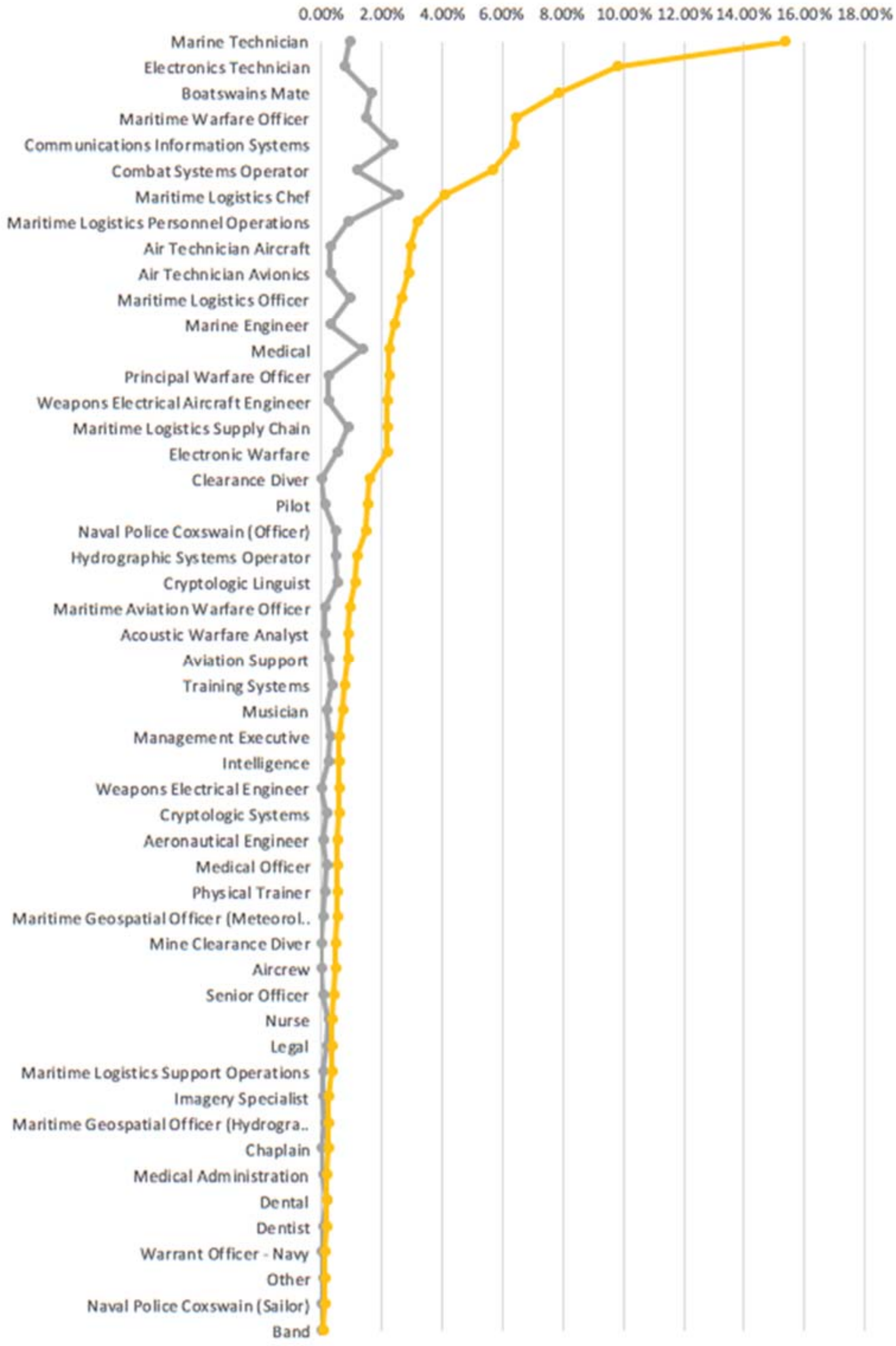


Figure 8. Gender density by function (workgroup).

E. THE "AVERAGE" RAN INDIVIDUAL

Overall, the average currently serving RAN individual is a male Sailor of E03 (Able Seaman) rank, aged between 25 and 29 years old whose functional employment category is 'Marine Technician.' These descriptive figures highlight the likely importance of gender differences in RAN retention and separation behavior.

F. METHODOLOGY

This thesis will produce length of service/survival profiles for the RAN Officers and Sailors with a focus on testing for heterogeneity in survival profiles along the dimensions of rank, workgroup, gender, age group and cohort year of enlistment. These survival profiles will inform the RAN what the probability is for a member remaining in the service at any point in their period of service.

After analyzing the characteristics variables (bivariate) individually and confirming their statistical significance, I will develop a preferred combined (multivariate) model specification based on these variables. This preferred model will then be analyzed and validated using Cox Proportional Hazard, logit and probit models to determine the best methodology for producing these length of service/survival profiles.

As described in detail in the preceding chapter, the Cox model estimates the Survival function, Hazard function and associated Hazard ratios.

Recall that a survival function, $S(t)$, is the probability of surviving past time t. It equals 1 at time $t=0$ and reduces as t goes to infinity. The survival function is described as:

$$S(t) = 1 - F(t) = Pr(T > t).$$

Also recall that the hazard function, $h(t)$, is the "probability that the failure event occurs in a given interval, conditional upon the subject having survived to the beginning of that interval" (Cleves et al., 2004, p. 7). The function is an "instantaneous rate of failure" (Cleves et al., 2004, p. 7) and is described as:

$$h(t) = \lim_{\Delta t \to 0} \frac{Pr(t+\Delta t > T > t | T > t)}{\Delta t} = \frac{f(t)}{S(t)}.$$

Finally, survival profiles will be produced for a few key variables to highlight their format and interpretation.

G. CHAPTER SUMMARY

This chapter has described the dataset and all of the variables. A descriptive statistics analysis has been conducted on the full RAN population snapshot so that we can understand the population and its characteristics. As of 31 May 2018 there were 13,727 Officers and Sailors serving in the RAN. The average RAN individual is a male Sailor of E03 (Able Seaman) rank, aged between 25 and 29 years old who's functional employment category is 'Marine Technician.'

The methodology for this thesis was also detailed to inform the next chapter (findings and results).

IV. FINDINGS AND RESULTS

In this chapter I will describe the findings and results from my multiple analyses. In Section A, I will investigate length of service survival profiles, including multiple and single spell specifications, Cox survival and hazard function curves and separation rates by length of service. In Section B, I will investigate the covariate characteristics of workgroup, rank, gender, calendar year of enlistment cohorts and age groups. From these available data set variables I will determine which account for RAN separation behavior. In Section C, I will conduct sensitivity analyses and validation techniques to confirm which model specification is valid. Finally, in Section D, I will summarize all of the findings and results.

A. SURVIVAL PROFILES/LOS

1. Multiple versus Single Spells

I address my primary research question: ‘What are the length of service/survival profiles for RAN Officers and Sailors?’, by estimating survival profiles of RAN service members. I begin by quantitatively determining the extent to which multiple service spells in the RAN will significantly impact model estimation. As discussed in previous chapters, the barriers between civilian, active duty, and reserve military service are fairly permeable. RAN policy allows for individuals to enlist, separate (or transfer to the reserves component) and re-enlist multiple times.

The RAN policy translates to a data structure where service members are likely to have multiple service spells, posing an additional complication to developing a survival model. Fortunately the Cox model is able to handle the multiple periods of service. In theory each spell may contain important information about the individuals, and so I will first consider modeling multiple spells and compare the results with a standard Cox model utilizing only the first spell.

To quantify the size of the single versus multiple spells problem, the total number of observations for each spell are presented in Table 9.

Table 9. Number of spells.

Number of spells	Freq.	Percent
1	21,173	97.03
2	638	2.92
3	9	0.04
Total	21,820	100

As can be seen in Table 9, by far the majority (97%) of observations in the dataset served (or are serving) a single spell (period of service) with 21,173 out of 21,820 observations. Next, 638 observations are for two spells (3%) and finally nine observations relate to three spells (less than 1%). From a simple statistics point of view there are not enough multiple spells to influence any analysis. However, I conducted a further check to determine the difference between a Cox regression with all multiple spells included and next only with the single spells. The output for these two regressions is in Table 10.

Table 10. Comparison table, Cox regressions for multiple spells and single spell.

	Multiple spell	Single spell
female	0.103*** (4.51)	0.101*** (4.44)
N	21495	21820

t statistics in parentheses
* p<0.05, **p<0.01, *** p<0.001

Note: In each column the top number represents the Cox regression coefficient, the middle number indicates the t-statistic and the bottom number reflects the number of observations relevant to this regression (n).

Table 10 shows a comparison of the outputs for two Cox regressions representing one for single spell data and a second for multiple spells, both regressed on the randomly selected variable 'female.' The coefficients can be interpreted as the difference between the baseline group 'male.' Both coefficients are very statistically significant, that the t-statistics are very similar and that the number of relevant observations has reduced from the multiple spell regression to the single spell regression. The fact that the coefficients and the t-statistics are so close means that the multiple spell setup is not necessarily required to add more to the analysis, however going forward I will use the multiple spell setup for consistency.

2. Separation Rates by LOS/cumLOS

The simplest way to determine unconditional survival or hazard rates is by calculating raw separation rates, as reported in Table 11.

The second row of Table 11 shows the separation rates (number of individuals who separate/total number of individuals in the sample) split across each cumulative length of service (cumLOS) in years. The main finding with this table is that the highest separation rate (10.3%) occurs at cumLOS=0, meaning that 10.3% of individuals separate from the RAN within their first year of service, without completing one year of service. The next highest separation rate of 6.6% occurs at cumLOS=6, meaning that 6.6% of all individuals separate having completed six years of service. This period aligns perfectly with the Initial Minimum Period of Service (IMPS) ‘contract period’ of 6 years for technical category Sailors.

The next highest separation rate of 4.9% occurs at cumLOS=4, meaning that 4.9% of individuals separate from the RAN after completing four years of service. This represents the non-technical category Sailors IMPS of four years

Table 11. Separation rates by cumulative length of service (in years).

	Cumulative Length of Service (in years)																	
Failed	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
(fail=0) Not Separate	3.5	4.0	4.2	4.3	5.0	3.7	2.7	3.0	3.7	3.1	2.6	2.5	2.0	1.6	1.7	1.6	0.5	49.8
(fail=1) Separate	10.3	4.9	3.4	3.1	4.9	3.4	6.6	3.4	3.2	1.9	1.9	1.4	0.9	0.6	0.3	0.2	0.0	50.2
Total	13.8	8.9	7.6	7.4	9.9	7.1	9.3	6.4	6.9	5.0	4.5	3.9	2.9	2.2	2.0	1.8	0.5	100.0

Note: Each cell represents a percentage.

Table 12 shows the separation rates by cumulative length of service in years for Sailors only. As can be seen the most significant separation rate is 11.2% and this again occurs at cumLOS=0, meaning that 11.2% of Sailors separate before completing one year of service. Like above the next two highest separation rates occur at cumLOS=6 and cumLOS=4.

Table 13 shows the separation rates by cumulative length of service in years for Officers only. As can be seen the most significant separation rate is 5.8% and this again occurs at cumLOS=0, meaning that 5.8% of Officers separate before completing one year of service. Unlike the Sailors in Table 12, the next two highest separation rates occur at cumLOS=1 and cumLOS=2. This shows the Officers separation pattern is quite different to the Sailors. In addition to this over the full 16 years of this data set, the Officer's separation rate is 38.6% whereas the Sailor's is 52.3% indicating that Officers are more likely to serve for longer periods than Sailors.

Table 12. Separation rates for Sailors only by cumulative length of service (in years).

Just Sailors	Cumulative Length of Service (in years)																	
Failed	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
(fail=0) Not Separate	3.3	3.6	4.0	4.2	5.1	3.8	2.6	2.8	3.6	3.1	2.6	2.3	1.9	1.4	1.6	1.5	0.5	47.7
(fail=1) Separate	11.2	4.8	3.3	3.1	5.2	3.5	7.2	3.7	3.3	2.0	2.0	1.4	0.8	0.5	0.3	0.2	0.0	52.3
Total	14.5	8.4	7.3	7.3	10.3	7.3	9.8	6.5	6.9	5.1	4.6	3.6	2.7	1.9	1.9	1.6	0.5	100.0

Note: Each cell represents a percentage.

Table 13. Separation rates for Officers only by cumulative length of service (in years).

Just Officers	Cumulative Length of Service (in years)																	
Failed	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
(fail=0) Not Separate	4.5	5.9	5.2	4.8	4.5	3.7	3.5	3.8	4.5	3.3	2.6	4.0	2.8	2.8	2.6	2.2	0.6	61.4
(fail=1) Separate	5.8	5.4	4.1	3.2	3.3	2.4	3.2	1.9	2.3	1.6	1.4	1.2	1.5	0.8	0.4	0.2	0.0	38.6
Total	10.3	11.4	9.3	8.1	7.7	6.1	6.7	5.7	6.8	4.9	4.0	5.2	4.3	3.6	3.0	2.5	0.6	100.0

Note: Each cell represents a percentage.

3. Cox Survival Function Curve

The survival (function) curve in Figure 9 displays the survival (function) rate (conditional probabilities) across all time periods.

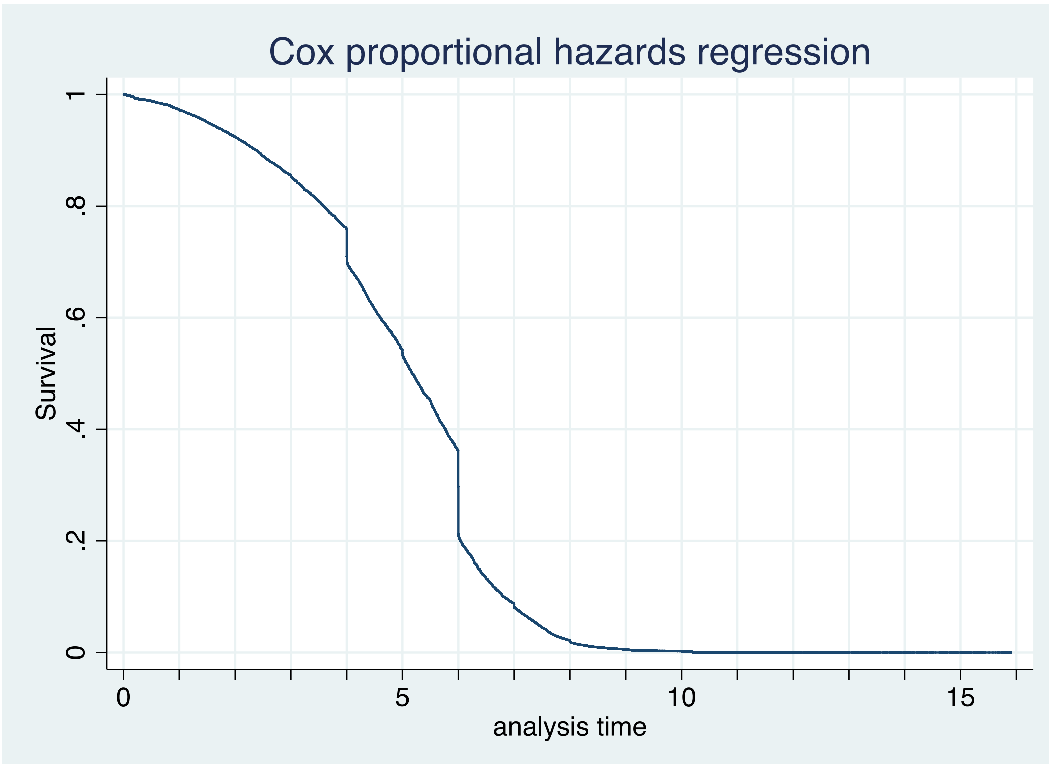


Figure 9. Cox survival (function) curve for the whole Navy.

Recall that a survival function, *S(t)*, is the probability of surviving past time *t*. In other words, "it is the probability that there is no failure event prior to *t*" (Cleves et al., 2004, p. 7). It equals 1 at time $t=0$ and reduces as *t* goes to infinity. It is also the reverse cumulative distribution function of T:

$$S(t) = 1 - F(t) = Pr(T > t).$$

As an example, to interpret this curve at six years of service: the probability of survival (remaining in the service) past six years of service is 0.20, this also can be interpreted as the probability that there is not a separation event prior to six years of service.

The shape of the survival curve indicates that initially there is a steady decline in the survival rate from year zero to year four. The steady decline is followed by a dramatic drop off at year four, then a steeper decline between years four and six, followed by a second more dramatic drop off at year six, and then finally a (parabolic) decline from years six onwards eventually approximating zero survival.

To summarize, both the raw separation rates and the survival curves (hazard rates) show that the most significant separation occurs during year zero, followed by year six, and then at year four.

B. CHARACTERISTICS

The next set of analyses addresses the question, 'What are the characteristics that could account for the observed survival profiles and any differences that are significant?'

This section attempts to determine the characteristics that are both statistically and economically significant in explaining RAN service members' separation behavior.

Interactions between two variables are used to help identify differential hazards by characteristics at each length of service (cumLOS). In layman's terms, an interaction between two variables allows the comparison of two groups of people holding constant something that they have in common (i.e., comparing females to males with the same length of service).

Given the above findings of significant variation in drop-offs in survival across years of service (e.g. LOS=6 is significantly different from LOS=1), the specifications below interact each of the key variables (workgroup, rank, gender, calendar year of enlistment cohorts and age groups) with cumulative length of service (cumLOS) interaction dummies.

1. Workgroup

a. Interaction between Aggregated Function at Enlistment (RevisedFunction0) and Cumulative Length of Service (cumLOS)

The first key explanatory variable I examine are the Workgroups (WGs) of RAN personnel. Table 14 reports the estimates from a survival model accounting for Workgroup indicators interacted with length of service.

Table 14. Regression output—Interaction between aggregated function at enlistment (RevisedFunction0) and cumulative length of service (cumLOS).

RevisedFunction0	**Interaction Variable**	**Hazard Ratio (s.e.)**	
Aeronautical Engineer	_IRevXcumL_2	1.326	***
		(0.090)	
Air Technician Aircraft	_IRevXcumL_3	1.380	***
		(0.081)	
Air Technician Avionics	_IRevXcumL_4	1.372	***
		(0.080)	
Aircrew	_IRevXcumL_5	1.458	***
		(0.124)	
Aviation Support	_IRevXcumL_6	1.009	
		(0.119)	
Boatswains Mate	_IRevXcumL_7	1.421	***
		(0.082)	
Chaplain	_IRevXcumL_8	0.955	
		(0.095)	
Clearance Diver	_IRevXcumL_9	1.390	***
		(0.082)	
Combat Systems Operator	_IRevXcumL_10	1.400	***
		(0.081)	
Combat Systems Operator Mine Warfare	_IRevXcumL_11	1.393	***
		(0.083)	
Communications Information Systems	_IRevXcumL_12	1.405	***
		(0.081)	
Communications Information Systems Submariner	_IRevXcumL_13	0.953	
		(0.116)	

RevisedFunction0	Interaction Variable	Hazard Ratio (s.e.)	
Cryptologic Linguist	_IRevXcumL_14	1.457	***
		(0.086)	
Cryptologic Systems	_IRevXcumL_15	1.419	***
		(0.086)	
Dental	_IRevXcumL_16	1.472	***
		(0.098)	
Dentist	_IRevXcumL_17	1.482	***
		(0.100)	
Electronic Warfare	_IRevXcumL_18	1.099	
		(0.087)	
Electronic Warfare Submarines	_IRevXcumL_19	1.209	*
		(0.091)	
Electronics Technician	_IRevXcumL_20	1.429	***
		(0.082)	
Electronics Technician Submariner	_IRevXcumL_21	1.165	*
		(0.082)	
General Experience	_IRevXcumL_22	1.308	***
		(0.081)	
Hydrographic Systems Operator	_IRevXcumL_23	1.399	***
		(0.084)	
Imagery Specialist	_IRevXcumL_24	1.392	*
		(0.202)	
Intelligence	_IRevXcumL_25	0.993	
		(0.197)	
Legal	_IRevXcumL_26	1.445	***
		(0.094)	
Management Executive	_IRevXcumL_27	0.884	
		(0.241)	
Marine Engineer	_IRevXcumL_28	1.221	***
		(0.076)	
Marine Engineer Submariner	_IRevXcumL_29	1.367	*
		(0.192)	
Marine Technician	_IRevXcumL_30	1.417	***
		(0.082)	
Marine Technician Submariner	_IRevXcumL_31	1.181	*
		(0.081)	

RevisedFunction0	Interaction Variable	Hazard Ratio (s.e.)	
Maritime Aviation Warfare Officer	_IRevXcumL_32	1.407	***
		(0.086)	
Maritime Geospatial Officer (Hydrographer)	_IRevXcumL_33	1.141	
		(0.189)	
Maritime Geospatial Officer (Meteorologist/Oceanographer)	_IRevXcumL_34	0.310	
		(19186.830)	
Maritime Logistics Chef	_IRevXcumL_35	1.427	***
		(0.083)	
Maritime Logistics Chef Submariner	_IRevXcumL_36	1.208	
		(0.140)	
Maritime Logistics Officer	_IRevXcumL_37	1.403	***
		(0.084)	
Maritime Logistics Personnel Operations	_IRevXcumL_38	1.400	***
		(0.082)	
Maritime Logistics Supply Chain	_IRevXcumL_39	1.307	***
		(0.077)	
Maritime Logistics Supply Chain Submariner	_IRevXcumL_40	0.000	
		(0.000)	
Maritime Logistics Support Operations	_IRevXcumL_41	1.386	***
		(0.081)	
Maritime Logistics Support Operations Submariner	_IRevXcumL_42	1.374	*
		(0.217)	
Maritime Warfare Officer	_IRevXcumL_43	1.364	***
		(0.079)	
Maritime Warfare Officer Submariner	_IRevXcumL_44	1.230	*
		(0.127)	
Medical	_IRevXcumL_45	1.389	***
		(0.082)	
Medical Administration	_IRevXcumL_46	1.285	
		(0.183)	
Medical Officer	_IRevXcumL_47	1.410	***
		(0.088)	
Medical Submariner	_IRevXcumL_48	0.546	
		(0.397)	
Mine Clearance Diver	_IRevXcumL_49	1.190	
		(0.195)	

RevisedFunction0	Interaction Variable	Hazard Ratio (s.e.)	
Musician	_IRevXcumL_50	1.322	***
		(0.083)	
Naval Police Coxswain (Officer)	_IRevXcumL_51	1.314	
		(0.243)	
Naval Police Coxswain (Sailor)	_IRevXcumL_52	1.415	***
		(0.087)	
Nurse	_IRevXcumL_53	0.696	***
		(0.060)	
Other Officers	_IRevXcumL_54	1.188	
		(0.266)	
Other Sailors	_IRevXcumL_55	1.406	***
		(0.083)	
Physical Trainer	_IRevXcumL_56	1.525	***
		(0.140)	
Pilot	_IRevXcumL_57	1.365	***
		(0.083)	
Principal Warfare Officer	_IRevXcumL_58	1.233	**
		(0.091)	
Training Systems	_IRevXcumL_59	1.418	***
		(0.088)	
Warrant Officer (Entry)	_IRevXcumL_60	1.396	***
		(0.099)	
Weapons Electrical Aircraft Engineer	_IRevXcumL_61	1.264	**
		(0.110)	
Weapons Electrical Engineer	_IRevXcumL_62	1.310	***
		(0.080)	
Weapons Electrical Engineer Submariner	_IRevXcumL_63	1.001	
		(0.228)	
	N	21820	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is ‘Acoustic Warfare Analyst’; Regression includes linear effect of cumLOS

The variable RevisedFunction0 represents an aggregated 'rolled-up' categorization of the lower level variable Function0. While Function0 contains more detailed information on the WGs, a separate set of regressions (not reported here, but available upon request) showed a lack of precision in the estimates. Around 87% of the coefficients are not statistically significant at the 0.05 level when categorizing personnel by Function0. The lack of precision in these estimates indicate that Function is perhaps not the best way to account for WG, as the cell sizes in each are too small. I thus re-categorized personnel based on Function0 into more aggregate classes.

Recall that the Cox proportional hazards model assumes that the hazard rate is a multiplicative proportion of the baseline hazard, where the multiplicative factor is an exponential function of the regressors. The baseline group in this regression is the first entry in the RevisedFunction0 variable, which is 'Acoustic Warfare Analyst'. When I estimate this regression I found that 74% of the coefficients are statistically significant at the 0.05 level or 95% degree of confidence. All of these also appear to be economically significant.

Interpreting these estimates, the Hazard Ratio (an odds ratio) for 'Marine Technician' of 1.417 indicates that they are 1.417 times as likely or 41.7% more likely to separate than 'Acoustic Warfare Analyst,' holding constant analysis time.

b. Interaction between Aggregated Function at separation (RevisedFunction1) and Cumulative Length of Service (cumLOS)

Next, I examine the Sailors' WG at separation or WG as of the snapshot on 31 May 2018. Table 15 includes the output of this regression.

Table 15. Regression output—Interaction between aggregated function at separation (RevisedFunction1) and cumulative length of service (cumLOS).

RevisedFunction1	**Interaction Variable**	**Hazard Ratio (s.e.)**	
Aeronautical Engineer	_IRevXcumL_2	1.099	*
		(0.052)	
Air Technician Aircraft	_IRevXcumL_3	1.141	***
		(0.039)	
Air Technician Avionics	_IRevXcumL_4	1.113	**
		(0.039)	
Aircrew	_IRevXcumL_5	1.178	***
		(0.050)	
Aviation Support	_IRevXcumL_6	0.979	
		(0.049)	
Band	_IRevXcumL_7	0.023	
		(27841.790)	
Boatswains Mate	_IRevXcumL_8	1.173	***
		(0.039)	
Chaplain	_IRevXcumL_9	0.790	**
		(0.069)	
Clearance Diver	_IRevXcumL_10	1.141	***
		(0.040)	
Combat Systems Operator	_IRevXcumL_11	1.200	***
		(0.040)	
Combat Systems Operator Mine Warfare	_IRevXcumL_12	1.126	**
		(0.042)	
Combat Systems Operator-Above	_IRevXcumL_13	0.000	
		(0.000)	
Combat Systems Operator-Above-ASAC	_IRevXcumL_14	0.000	
		(274.731)	
Combat Systems Operator-Under	_IRevXcumL_15	0.000	
		(0.000)	
Communications Information Systems	_IRevXcumL_16	1.165	***
		(0.039)	

RevisedFunction1	Interaction Variable	Hazard Ratio (s.e.)	
Communications Information Systems Submariner	_IRevXcumL_17	1.045	
		(0.051)	
Cryptologic Linguist	_IRevXcumL_18	1.187	***
		(0.042)	
Cryptologic Systems	_IRevXcumL_19	1.174	***
		(0.048)	
Dental	_IRevXcumL_20	1.211	***
		(0.053)	
Dentist	_IRevXcumL_21	1.214	***
		(0.058)	
Electronic Warfare	_IRevXcumL_22	1.101	*
		(0.041)	
Electronic Warfare Submarines	_IRevXcumL_23	1.069	
		(0.049)	
Electronics Technician	_IRevXcumL_24	1.171	***
		(0.039)	
Electronics Technician Submariner	_IRevXcumL_25	1.047	
		(0.044)	
General Experience	_IRevXcumL_26	2.576	***
		(0.271)	
Hydrographic Systems Operator	_IRevXcumL_27	1.164	***
		(0.042)	
Imagery Specialist	_IRevXcumL_28	1.176	**
		(0.060)	
Intelligence	_IRevXcumL_29	1.018	
		(0.060)	
Legal	_IRevXcumL_30	1.193	***
		(0.052)	
Management Executive	_IRevXcumL_31	1.088	
		(0.060)	
Marine Engineer	_IRevXcumL_32	1.038	
		(0.041)	
Marine Engineer Submariner	_IRevXcumL_33	1.090	
		(0.076)	
Marine Technician	_IRevXcumL_34	1.174	***

RevisedFunction1	Interaction Variable	Hazard Ratio (s.e.)	
		(0.039)	
Marine Technician Submariner	_IRevXcumL_35	1.112	**
		(0.041)	
Maritime Aviation Warfare Officer	_IRevXcumL_36	1.136	**
		(0.047)	
Maritime Geospatial Officer (Hydrographer)	_IRevXcumL_37	1.171	***
		(0.047)	
Maritime Geospatial Officer (Meteorologist/Oceanographer)	_IRevXcumL_38	1.215	***
		(0.057)	
Maritime Logistics Chef	_IRevXcumL_39	1.187	***
		(0.040)	
Maritime Logistics Chef Submariner	_IRevXcumL_40	1.123	*
		(0.064)	
Maritime Logistics Officer	_IRevXcumL_41	1.143	***
		(0.041)	
Maritime Logistics Personnel Operations	_IRevXcumL_42	1.163	***
		(0.040)	
Maritime Logistics Supply Chain	_IRevXcumL_43	1.102	**
		(0.038)	
Maritime Logistics Supply Chain Submariner	_IRevXcumL_44	0.850	
		(0.132)	
Maritime Logistics Support Operations	_IRevXcumL_45	1.159	***
		(0.040)	
Maritime Logistics Support Operations Submariner	_IRevXcumL_46	1.170	**
		(0.068)	
Maritime Warfare Officer	_IRevXcumL_47	1.123	***
		(0.038)	
Maritime Warfare Officer Submariner	_IRevXcumL_48	1.091	*
		(0.047)	
Medical	_IRevXcumL_49	1.149	***
		(0.040)	
Medical Administration	_IRevXcumL_50	1.106	
		(0.097)	
Medical Officer	_IRevXcumL_51	1.160	***
		(0.047)	

RevisedFunction1	**Interaction Variable**	**Hazard Ratio (s.e.)**	
Medical Submariner	_IRevXcumL_52	0.999	
		(0.106)	
Mine Clearance Diver	_IRevXcumL_53	1.122	*
		(0.059)	
Musician	_IRevXcumL_54	1.096	*
		(0.045)	
Naval Police Coxswain (Officer)	_IRevXcumL_55	1.153	*
		(0.081)	
Naval Police Coxswain (Sailor)	_IRevXcumL_56	1.134	***
		(0.040)	
Nurse	_IRevXcumL_57	0.598	***
		(0.042)	
Other Officers	_IRevXcumL_58	1.312	***
		(0.087)	
Other Sailors	_IRevXcumL_59	2.980	***
		(0.470)	
Physical Trainer	_IRevXcumL_60	1.153	***
		(0.048)	
Pilot	_IRevXcumL_61	1.127	**
		(0.043)	
Principal Warfare Officer	_IRevXcumL_62	1.068	
		(0.044)	
Principal Warfare Officer Amphib	_IRevXcumL_63	0.001	
		(12291.310)	
Senior Officer	_IRevXcumL_64	1.131	
		(0.212)	
Training Systems	_IRevXcumL_65	1.141	***
		(0.046)	
Warrant Officer (Entry)	_IRevXcumL_66	1.225	***
		(0.071)	
Weapons Electrical Aircraft Engineer	_IRevXcumL_67	0.986	
		(0.085)	
Weapons Electrical Engineer	_IRevXcumL_68	1.079	*
		(0.042)	
Weapons Electrical Engineer Submariner	_IRevXcumL_69	1.156	**
		(0.064)	

RevisedFunction1	**Interaction Variable**	**Hazard Ratio (s.e.)**
	N	21820

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is ‘Acoustic Warfare Analyst’; Regression includes linear effect of cumLOS

The baseline group is the same WG as above. This regression indicates 72% of the coefficients are statistically significant at the 0.05 level. All of these also appear to be economically significant.

Tables 14 and 15 jointly indicate the significance of WGs in accounting for separation behavior of RAN personnel.

2. Rank

a. Interaction between Rank at Enlistment (RankCode0) and Cumulative Length of Service (cumLOS)

I next examine the extent to which Rank at enlistment and at separation (or current status, if not separated) significantly account for the estimated survival profiles. Table 16 contains the output for this regression.

Table 16. Regression output—Interaction between rank at enlistment (RankCode0) and cumulative length of service (cumLOS).

RankCode0	Interaction Variable	Hazard Ratio (s.e.)	
E01	_IRanXcumL_2	0.945	***
		(0.016)	
E02	_IRanXcumL_3	0.952	*
		(0.019)	
E03	_IRanXcumL_4	0.958	***
		(0.011)	
E05	_IRanXcumL_5	0.974	*
		(0.012)	
E06	_IRanXcumL_6	0.947	***
		(0.015)	
E08	_IRanXcumL_7	0.972	
		(0.018)	
E09	_IRanXcumL_8	0.713	
		(0.200)	
O00	_IRanXcumL_9	0.962	***
		(0.006)	
O01	_IRanXcumL_10	0.927	
		(0.078)	
O02	_IRanXcumL_11	0.966	**
		(0.013)	
O03	_IRanXcumL_12	0.879	***
		(0.016)	
O04	_IRanXcumL_13	0.902	***
		(0.022)	
O05	_IRanXcumL_14	0.941	
		(0.041)	
O06	_IRanXcumL_15	0.749	
		(0.172)	
	N	21820	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Baseline group is ‘E00’; Regression includes linear effect of cumLOS

I begin with a model lumping all ranks together.

The baseline group for this regression is the rank code 'E00' representing the initial Sailor rank of 'Recruit.' 64% of the coefficients in this regression are statistically significant at the 0.05 level.

This regression, however, includes all Sailor and Officer ranks and compares them to the initial Sailor rank code E00. This is deemed as not relevant for comparison purposes and the two groups should be separated and reanalyzed.

b. Interaction between Sailor Rank at Enlistment (SRank0) and Cumulative Length of Service (cumLOS)

I next estimate models separately accounting for enlisted Sailor ranks at enlistment and separation. The baseline is the same as above; however, this regression only includes the Sailor ranks. Table 17 displays the regression output.

Table 17. Regression output—Interaction between Sailor rank at enlistment (SRank0) and cumulative length of service (cumLOS).

SRank0	Interaction Variable	Hazard Ratio (s.e.)	
E01	_ISRaXcumLO_2	0.943	***
		(0.016)	
E02	_ISRaXcumLO_3	0.953	*
		(0.019)	
E03	_ISRaXcumLO_4	0.958	***
		(0.011)	
E05	_ISRaXcumLO_5	0.975	*
		(0.012)	
E06	_ISRaXcumLO_6	0.947	***
		(0.015)	
E08	_ISRaXcumLO_7	0.973	
		(0.018)	
E09	_ISRaXcumLO_8	0.715	
		(0.200)	
	N	18265	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Baseline group is ‘E00’; Regression includes linear effect of cumLOS

71% of the coefficients in this regression are statistically significant at the 0.05 level.

The Hazard Ratio for Sailor rank code ‘E03’ of 0.958 indicates that they are 0.96 times as likely or 4.2% less likely to separate if their rank at enlistment was Able Seaman (‘E03’) compared to those who joined as Sailor rank Recruit (‘E00’), holding constant analysis time. This also makes logical sense in that all enlistees with no prior military experience will join as a Recruit (E00) and face a higher separation rate due to the natural matching losses, whereas those who join as an Able Seaman (E03) are those who do have prior military experience, are on their 2nd or higher period of enlistment, are probably a little older, and are therefore less effected by the matching loss.

c. Interaction between Sailor Rank at Separation (SRank1) and Cumulative Length of Service (cumLOS)

I next estimate a model accounting for Sailor ranks at separation. The baseline group is the same as above. The regression outputs are displayed in Table 18.

Table 18. Regression output—Interaction between Sailor rank at separation (SRank1) and cumulative length of service (cumLOS).

SRank1	**Interaction Variable**	**Hazard Ratio (s.e.)**
E01	_ISRaXcumLO_2	0.438 (0.013)
E02	_ISRaXcumLO_3	0.384 (0.010)
E03	_ISRaXcumLO_4	0.416 (0.008)
E05	_ISRaXcumLO_5	0.424 (0.008)
E06	_ISRaXcumLO_6	0.431 (0.008)
E08	_ISRaXcumLO_7	0.406 (0.010)
E09	_ISRaXcumLO_8	0.383 (0.021)
	N	18265

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is 'E00'; Regression includes linear effect of cumLOS

100% of the coefficients in this regression are statistically significant at the 0.05 level.

The Hazard Ratio for Sailor rank code 'E05' of 0.424 indicates that they are 0.424 times as likely or 57.6% less likely to separate if their rank at separation was Leading Seaman ('E05') compared to those who separated as a Recruit ('E00'), holding constant analysis time. This does appear valid because those who separate at E00 didn't get promoted, holding constant analysis time.

Tables 17 and 18 jointly indicate the significance of Sailor ranks at enlistment and separation in accounting for separation behavior of RAN personnel.

d. Interaction between Officer Rank at Enlistment (ORank0) and Cumulative Length of Service (cumLOS)

I next estimate models separately accounting for Officer ranks at enlistment and separation. The baseline for this regression is the rank code 'O00' which represents the initial Officer rank 'Midshipman.' Table 19 contains the output.

Table 19. Regression output—Interaction between Officer rank at enlistment (ORank0) and cumulative length of service (cumLOS).

ORank0	**Interaction Variable**	**Hazard Ratio (s.e.)**	
O01	_IORaXcumLO_2	0.968	
		(0.081)	
O02	_IORaXcumLO_3	1.003	
		(0.014)	
O03	_IORaXcumLO_4	0.959	*
		(0.017)	
O04	_IORaXcumLO_5	0.938	*
		(0.024)	
O05	_IORaXcumLO_6	0.983	
		(0.043)	
O06	_IORaXcumLO_7	0.796	
		(0.179)	
	N	3380	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is 'O00'; Regression includes linear effect of cumLOS

Of the coefficients in this regression, 67% are not statistically significant at the 0.05 level.

The Hazard Ratio for Officer rank code ‘O03’ of 0.959 indicates that they are 0.959 times as likely or 4.1% less likely to separate if their rank at enlistment was Lieutenant (‘O03’) compared to those who joined as Officer rank Midshipman (‘O00’), holding constant analysis time. Like the Sailors rank at enlistment situation above, this also makes logical sense in that all Officer enlistees with no prior military experience will join as a Midshipman (O00) and face a higher separation rate due to the natural matching losses, whereas those who join as an Lieutenant (O03) are those who do have prior military experience, are on their 2nd or higher period of enlistment, are probably a little older, or have been recruited at this rank because of their specialist skills, and are therefore less effected by the matching loss.

Due to the smaller proportion of statistical significance this variable is not suitable to include in the preferred model, meaning that the rank at enlistment for Officers does not significantly contribute to separation behavior of RAN personnel.

e. Interaction between Officer Rank at Separation (ORank1) and Cumulative Length of Service (cumLOS)

I next estimate a model accounting for Officer ranks at separation. The baseline group is the same as above. Table 20 contains the regression output.

Table 20. Regression output—Interaction between Officer rank at separation (ORank1) and cumulative length of service (cumLOS).

ORank1	Interaction Variable	Hazard Ratio (s.e.)	
O01	_IORaXcumLO_2	0.974	
		(0.033)	
O02	_IORaXcumLO_3	0.892	***
		(0.023)	
O03	_IORaXcumLO_4	0.894	***
		(0.020)	
O04	_IORaXcumLO_5	0.895	***
		(0.022)	
O05	_IORaXcumLO_6	0.874	***
		(0.030)	
O06	_IORaXcumLO_7	0.751	*
		(0.108)	
O07	_IORaXcumLO_8	0.950	
		(0.174)	
	N	3555	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is ‘O00’; Regression includes linear effect of cumLOS

71% of the coefficients in this regression are statistically significant at the 0.05 level.

The Hazard Ratio for Officer rank code ‘O04’ of 0.895 indicates that they are 0.895 times as likely or 10.5% less likely to separate if their rank at separation was Lieutenant Commander (‘O04’) compared to those who separated as a Midshipman (‘O00’), holding constant analysis time.

Table 20 indicates the significance of Officer rank at separation in accounting for separation behavior of RAN personnel.

3. Gender

a. Bivariate Model with Female Only

To investigate whether survival profiles significantly vary by gender of RAN personnel, I first estimate a survival model controlling for just gender. Table 21 shows that gender is indeed significant, with the female coefficient statistically significantly different from 0. Females are 1.1 times or 10% more likely to separate than males across the analysis time in the study.

Table 21. Bivariate model with female only.

Gender	**Interaction Variable**	**Hazard Ratio (s.e.)**	
female	N/a	1.107 (0.025)	***
	N	5040	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Baseline group is 'male'

b. Interaction between Female and cumLOS

However, when I include interactions of gender with length of service, gender by itself is no longer statistically significant. Table 22 reports estimates from a survival model accounting for gender indicators interacted with length of service. This regression shows that gender interacted with length of service is not statistically significant.

Table 22. Interaction between female and cumLOS.

Gender with cumLOS	**Interaction Variable**	**Hazard Ratio (s.e.)**
female (=1)	_IfemXcumLO_1	1.002 (0.004)
	N	5040

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is 'male'

c. Interaction between Female and cumLOS (PlusFemale)

Next, I estimated a model with linear and interacted female and cumLOS. Table 23 shows that the included gender indicators were also not statistically significant.

Table 23. Interaction between female and cumLOS (plus female).

Gender	**Interaction Variable**	**Hazard Ratio (s.e.)**
female	N/A	0.947
		(0.036)
female (=1)	_IfemXcumLO_1	1.011
		(0.007)
	N	5040

Note: Exponentiated coefficients; Standard errors in parentheses
Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001
Baseline group is 'male'

d. Dummies between i.female and i.cumLOS

One hypothesis is that there are possible nonlinearities in how gender accounts for survival profiles among RAN personnel. One way to get at this is by estimating a model with full interactions of female and each value of cumulative length of service. This specification would then be able to determine whether there were differing gender effects occurring amongst the personnel with the same cumulative length of service. Table 24 displays the regression output.

Table 24 reports some interesting results. Fifty percent of the coefficients were statistically significant. More importantly, the interactions of female indicator with length of service are statistically significantly different from zero from years one to nine (except at year six), and not statistically different at length of service 10 years and onward. In other words, gender differences are particularly more salient among those with less than 10 years of service.

Table 24. Regression output—Dummies between i.female and i.cumLOS.

Gender and cumLOS	Interaction Variable	Hazard Ratio (s.e.)	
female=1,cumLOS=1	_IcumXfem_1_1	0.683	***
		(0.062)	
female=1,cumLOS=2	_IcumXfem_2_1	1.388	***
		(0.141)	
female=1,cumLOS=3	_IcumXfem_3_1	1.824	***
		(0.184)	
female=1,cumLOS=4	_IcumXfem_4_1	1.644	***
		(0.142)	
female=1,cumLOS=5	_IcumXfem_5_1	1.406	***
		(0.137)	
female=1,cumLOS=6	_IcumXfem_6_1	1.142	
		(0.098)	
female=1,cumLOS=7	_IcumXfem_7_1	1.290	*
		(0.135)	
female=1,cumLOS=8	_IcumXfem_8_1	1.241	*
		(0.134)	
female=1,cumLOS=9	_IcumXfem_9_1	1.526	***
		(0.188)	
female=1,cumLOS=10	_IcumXfem_10_1	0.966	
		(0.134)	
female=1,cumLOS=11	_IcumXfem_11_1	1.288	
		(0.199)	
female=1,cumLOS=12	_IcumXfem_12_1	1.228	
		(0.218)	
female=1,cumLOS=13	_IcumXfem_13_1	1.211	
		(0.286)	
female=1,cumLOS=14	_IcumXfem_14_1	1.308	
		(0.395)	

Gender and cumLOS	Interaction Variable	Hazard Ratio (s.e.)
female=1,cumLOS=15	_IcumXfem_15_1	1.350
		(0.574)
female=1,cumLOS=16	_IcumXfem_16_1	6.410
		(9.070)
	N	21820

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Other regressors include i.cumLOS and i.female, however were excluded from the output table for formatting purposes.

The Hazard Ratio for female interacted with cumLOS=1 (_IcumXfem_1_1) of 0.6825622 indicates that females are 0.68 times as likely or 31.7% less likely to separate than males with cumLOS=1.

Between cumLOS of two to nine years, females are significantly more likely than males to separate. The Hazard Ratios for females with between two and nine cumLOS shows that females are between 14–82% more likely to separate than men with the same cumLOS. All Hazard Ratios after cumLOS=9 are not statistically significant, meaning that they are not statistically different from 0, or that females in these groups are not different from men in the same groups.

The difference in both statistical and economic significance of how gender affects separation in the RAN across early cumLOS and more than 10 years of cumLOS could be causing an averaging effect in the earlier specifications shown in the Tables above. That gender matters more in the earlier years could also coincide with gender differences across the life cycle, as these earlier years coincide with prime childbearing and childrearing years.

Results from Tables 21–24 indicate the significance of gender in accounting for separation behavior of RAN personnel, particularly in the early years.

4. Entry Cohorts (Calendar Year of Enlistment)

a. Interaction between Cohort Year of Enlistment (year) and Cumulative Length of Service (cumLOS)

The next key explanatory variable I examine are the entry cohorts (calendar year of enlistment) of RAN personnel. This model is included to determine whether cohort and/or peer effects are relevant to the separation behavior of RAN personnel. Table 25 reports the estimates from a survival model accounting for cohort year of enlistment indicators interacted with length of service.

Table 25. Regression output—Interaction between cohort year of enlistment (year) and cumulative length of service (cumLOS).

Calendar year of entry Cohort	**Interaction Variable**	**Hazard Ratio (s.e.)**	
2003	_IyeaXcum_2003	0.964	***
		(0.007)	
2004	_IyeaXcum_2004	0.951	***
		(0.007)	
2005	_IyeaXcum_2005	0.938	***
		(0.008)	
2006	_IyeaXcum_2006	0.914	***
		(0.007)	
2007	_IyeaXcum_2007	0.894	***
		(0.007)	
2008	_IyeaXcum_2008	0.874	***
		(0.008)	
2009	_IyeaXcum_2009	0.840	***
		(0.007)	
2010	_IyeaXcum_2010	0.760	***
		(0.008)	
2011	_IyeaXcum_2011	0.766	***
		(0.008)	
2012	_IyeaXcum_2012	0.690	***
		(0.011)	
2013	_IyeaXcum_2013	0.613	***
		(0.011)	

Calendar year of entry Cohort	**Interaction Variable**	**Hazard Ratio (s.e.)**	
2014	_IyeaXcum_2014	0.392	***
		(0.010)	
2015	_IyeaXcum_2015	0.242	***
		(0.010)	
2016	_IyeaXcum_2016	0.073	***
		(0.006)	
2017	_IyeaXcum_2017	0.001	***
		(0.000)	
2018	_IyeaXcum_2018	0.000	
		(0.000)	
	N	21820	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is calendar year of entry cohort ‘2002’; Regression includes linear effect of cumLOS

The baseline group for this regression is the calendar year of enlistment cohort 2002. Ninety-four percent of the coefficients in this regression are statistically significant.

The only cohort year which was not statistically significant is year 2018 which appears valid because it is not a full calendar year as the data for this year only spans from 1 Jan to 31 May 18.

It appears that there are potentially two factors at play here: a mechanical one, in that the earlier the cohort year of entry the higher the cumulative length of service and therefore the more likely to separate (as can be seen from the steady reduction in hazard ratios as year of entry increases); the second is that of the economic conditions in effect during the cohort entry year which may affect the outcome (separation probability).

b. *Interaction between Dummies for Each Cohort Year of Entry (Year) and Dummies for Each Cumulative Length of Service (cumLOS)*

To investigate the latter effect further, I estimate a model accounting for dummies for each cohort year of entry and dummies for each cumLOS to determine whether there are two effects. To implement this, I made the baseline group the cohort year of entry 2008 which aligns with the onset of the Global Financial Crisis (GFC) in Australia. By doing this I am attempting to compare all other cohort years to what was economically a significant year and determine whether a significant economic event would affect survival behavior of RAN personnel. Table 26 displays the output of the regression.

Table 26. Regression output—Interaction between dummies for each cohort year of entry (year) and dummies for each cumulative length of service (cumLOS).

Calendar year of entry Cohort (2008 baseline)	**Interaction Variable**	**Hazard Ratio (s.e.)**	
2002	_IyeaXcum_2002	1.145	***
		(0.010)	
2003	_IyeaXcum_2003	1.104	***
		(0.008)	
2004	_IyeaXcum_2004	1.089	***
		(0.008)	
2005	_IyeaXcum_2005	1.074	***
		(0.008)	
2006	_IyeaXcum_2006	1.046	***
		(0.008)	
2007	_IyeaXcum_2007	1.023	**
		(0.008)	
2009	_IyeaXcum_2009	0.961	***
		(0.008)	
2010	_IyeaXcum_2010	0.870	***
		(0.009)	
2011	_IyeaXcum_2011	0.877	***
		(0.009)	
2012	_IyeaXcum_2012	0.789	***

Calendar year of entry Cohort (2008 baseline)	**Interaction Variable**	**Hazard Ratio (s.e.)**	
		(0.012)	
2013	_IyeaXcum_2013	0.702	***
		(0.012)	
2014	_IyeaXcum_2014	0.448	***
		(0.011)	
2015	_IyeaXcum_2015	0.277	***
		(0.011)	
2016	_IyeaXcum_2016	0.084	***
		(0.007)	
2017	_IyeaXcum_2017	0.001	***
		(0.000)	
2018	_IyeaXcum_2018	0.000	
		(0.000)	
	N	21820	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is calendar year of entry cohort '2008'

Like the last regression 94% of Hazard Ratios in this regression are statistically significant. However, the Hazard Ratios on the years 2002–2007 (1.144–1.023) indicate that personnel joining in these years are more likely to separate compared to those who joined in 2008. For those whose cohort year is after 2008 (i.e. 2009 onwards) the Hazard Ratios (0.961–0.001) indicates that personnel joining after 2008 are less likely to separate compared to those who joined in 2008. This outcome can be interpreted that the GFC in Australia may have led to a lower separation rate for all who joined after 2008.

Table 26 indicates the significance of cohort year of enlistment in accounting for separation behavior of RAN personnel, and that cohort year may proxy for macroeconomic conditions.

5. Age Groups

a. Interaction between Age Groups at Enlistment (AG0) and Cumulative Length of Service (cumLOS)

The next key explanatory variable I examine are the age groups at enlistment and separation of RAN personnel interacted with cumLOS. Table 27 contains the regression output. Here this interaction between age groups at enlistment and with cumLOS attempts to determine whether the age of an individual at enlistment affects their separation rate for those individuals with the same cumLOS.

Table 27. Regression output—Interaction between age groups at enlistment (AG0) and cumulative length of service (cumLOS).

Age Group 0	**Interaction Variable**	**Hazard Ratio (s.e.)**	
Age2024	_IAG0XcumLO_2	0.983	***
		(0.004)	
Age2529	_IAG0XcumLO_3	0.923	***
		(0.006)	
Age3034	_IAG0XcumLO_4	0.947	***
		(0.008)	
Age3539	_IAG0XcumLO_5	0.957	***
		(0.010)	
Age4044	_IAG0XcumLO_6	0.944	***
		(0.011)	
Age4549	_IAG0XcumLO_7	0.947	***
		(0.013)	
Age5054	_IAG0XcumLO_8	0.933	**
		(0.023)	
Age5559	_IAG0XcumLO_9	1.053	
		(0.054)	
	N	21820	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is age group at enlistment 'Age1619'; Regression includes linear effect of cumLOS

The baseline group for this regression was the 16−19 years old category. All age groups, except for one—the 55−59 years old category, are statistically significantly different from zero.

While statistically significant, the magnitudes of the Hazard Ratios in Table 27 indicate all of them are not economically significant. The estimated hazard ratios range from 0.92 to 0.98 (meaning minimal effect on separation).

b. Interaction between Age Groups at Separation (AG1) and Cumulative Length of Service (cumLOS)

I next estimate a model accounting for age group at separation interacted with cumLOS and Table 28 contains the output.

Table 28. Regression output—Interaction between age groups at separation (AG1) and cumulative length of service (cumLOS).

Age Group 1	Interaction Variable	Hazard Ratio (s.e.)	
Age2024	_IAG1XcumL_2	0.579	***
		(0.015)	
Age2529	_IAG1XcumL_3	0.555	***
		(0.014)	
Age3034	_IAG1XcumL_4	0.563	***
		(0.015)	
Age3539	_IAG1XcumL_5	0.559	***
		(0.015)	
Age4044	_IAG1XcumL_6	0.559	***
		(0.015)	
Age4549	_IAG1XcumL_7	0.545	***
		(0.015)	
Age5054	_IAG1XcumL_8	0.547	***
		(0.016)	
Age5559	_IAG1XcumL_9	0.567	***
		(0.017)	
Age6064	_IAG1XcumL_10	0.589	***
		(0.020)	
	N	21820	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is calendar year of entry cohort 'Age1619'; Regression includes linear effect of cumLOS

The baseline group is the same as above. 100% of the coefficients in this regression are statistically significant. Older age groups are less likely to separate than the baseline group of aged 16−19 years old at separation. The Hazard Ratios are difficult to interpret in this specification, however, as cumLOS and age could be linearly dependent.

c. Regression of Age Groups at Separation (AG1) Only

Because of the linear relationship between cumLOS and age, I next estimate a model accounting for age group at separation only (without the cumLOS interactions). Table 29 displays the regression output.

Table 29. Regression output—Age groups at separation (AG1) only.

Age Group 1	**Variable**	**Hazard Ratio (s.e.)**	
Age2024	_IAG1_2	0.140	***
		(0.005)	
Age2529	_IAG1_3	0.035	***
		(0.001)	
Age3034	_IAG1_4	0.014	***
		(0.001)	
Age3539	_IAG1_5	0.015	***
		(0.001)	
Age4044	_IAG1_6	0.018	***
		(0.001)	
Age4549	_IAG1_7	0.019	***
		(0.001)	
Age5054	_IAG1_8	0.017	***
		(0.002)	
Age5559	_IAG1_9	0.013	***
		(0.002)	
Age6064	_IAG1_10	0.021	***
		(0.004)	
	N	21820	

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline group is age group at separation 'Age1619'

The baseline group was set as those in their late 20's. The upshot is that these specifications revealed those who enlisted before age 20 are such outliers relative to all the other age groups (meaning the Hazard ratios for the youngest age groups are the highest). Among those who enlist above age 20, the differences across, e.g., age 20–24 versus 25–29 versus 30–34 appear to be smooth and differences in hazard rates are as expected. Those who enlist in their late 50s close to retirement age also have hazard ratios that are significantly different, but not quite as dramatically different as those who enlist prior to 20.

Table 29 indicates the significance of age at separation in accounting for separation behavior of RAN personnel.

d. *Mean Failure Rate and Mean Cumulative Length of Service by Age Group at Enlistment*

An alternative way to view the same data is by calculating mean separation rates, as reported in Table 30.

Table 30. Mean failure and mean cumulative length of service by each age group at enlistment.

Age Group at enlistment	Mean failure (separation) rate	Mean cumulative length of service (in years)
16–19	0.539809	5.324521
20–24	0.491767	5.236018
25–29	0.439788	5.340521
30–34	0.467206	5.742138
35–39	0.466877	5.951104
40–44	0.401361	6.207483
45–49	0.457064	5.401662
50–54	0.429577	4.422535
55–59	0.5	4.083333

Table 30 displays the mean failure (separation) rate (in column 1) along with the mean cumulative length of service (column 2) by each age group at enlistment. What is significant here is that the mean failure (separation) rate is highest for the youngest age group 16–19, and is lower for all other age groups. This means that at a point in time the younger age groups are more likely to separate from the RAN than the older age groups.

In addition to this, the right hand column displays the mean cumulative length of service for each age group. This column shows that the average length of service is approximately 5.2–5.3 years for individuals under 30, however for individuals over 30 years the length of service increases to 5.7–6.2 years.

e. *Mean Failure (Separation) Rate by Age Group at Enlistment and Year of Entry*

Yet another alternative way to view the same data is by calculating mean separation rates by cohort year of enlistment and age group at enlistment, as reported in Table 31.

Table 31. Mean failure (separation) rate by age group at enlistment and year of entry.

	Age Group at enlistment (AG0)									
Year of Entry	**16-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**Total**
2002	265 0.85 6.82	187 0.84 7.37	99 0.78 8.03	45 0.80 8.60	20 0.80 9.75	11 0.73 10.00	12 1.00 8.50	1 1.00 3.00		640 0.83 7.47
2003	799 0.79 7.39	443 0.81 7.42	187 0.73 8.43	90 0.76 8.00	46 0.80 7.80	31 0.61 9.39	17 0.94 6.65	3 1.00 3.00		1616 0.79 7.59
2004	788 0.79 7.28	475 0.74 7.80	139 0.76 7.12	89 0.75 7.37	36 0.78 6.58	40 0.68 8.33	11 0.91 7.36	4 1.00 4.75		1582 0.77 7.43
2005	576 0.72 7.65	349 0.74 7.29	113 0.73 7.23	58 0.76 7.09	43 0.65 8.77	35 0.60 9.11	13 0.69 9.54	2 0.50 8.50	1 1.00 8.00	1190 0.72 7.58
2006	714	424	125	42	39	37	15	8	1	1405

	Age Group at enlistment (AG0)									
Year of Entry	**16-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**Total**
	0.72	0.70	0.68	0.57	0.54	0.68	0.60	0.63	1.00	0.70
	6.87	7.38	7.38	8.36	7.92	8.11	9.07	8.00	6.00	7.20
2007	759	500	149	83	58	35	25	4		1613
	0.70	0.68	0.62	0.69	0.55	0.54	0.60	0.75		0.67
	6.79	6.92	7.15	6.76	7.91	7.80	8.32	7.75		6.95
2008	803	449	128	48	48	38	21	7	2	1544
	0.65	0.63	0.64	0.65	0.56	0.47	0.62	0.86	1.00	0.64
	6.07	6.33	6.77	6.81	6.71	7.39	6.95	4.86	3.50	6.29
2009	796	501	161	87	52	53	38	15	1	1704
	0.65	0.61	0.54	0.51	0.60	0.47	0.58	0.53	1.00	0.61
	5.74	6.09	6.35	6.67	6.02	7.08	6.47	7.27	3.00	6.03
2010	654	531	191	90	62	60	36	14	3	1641
	0.54	0.51	0.49	0.48	0.40	0.43	0.56	0.71	1.00	0.51
	5.59	5.87	6.42	6.03	6.37	6.12	5.39	5.79	3.67	5.84
2011	549	408	131	58	35	38	34	11	1	1265
	0.51	0.50	0.46	0.40	0.54	0.45	0.38	0.55	1.00	0.49
	5.20	5.36	5.69	5.62	5.11	5.76	6.00	5.09	7.00	5.36
2012	396	282	109	41	29	31	13	6		907
	0.35	0.35	0.39	0.27	0.34	0.29	0.23	0.50		0.35
	4.95	5.00	5.06	5.37	4.83	5.42	5.69	4.83		5.02
2013	488	369	129	75	24	54	26	12	4	1181
	0.34	0.31	0.23	0.31	0.21	0.09	0.23	0.33	0.75	0.30
	3.95	4.13	4.67	4.96	4.63	4.78	5.04	5.00	6.25	4.24
2014	596	481	204	70	33	47	25	10		1466
	0.28	0.22	0.20	0.29	0.21	0.19	0.28	0.20		0.24
	3.28	3.48	3.92	3.93	3.88	3.66	3.36	4.30		3.50
2015	467	397	174	81	27	25	15	10	1	1197
	0.21	0.21	0.16	0.20	0.15	0.20	0.20	0.10	0.00	0.20
	2.54	2.61	3.09	3.72	3.11	3.00	3.20	2.80	3.00	2.76
2016	398	360	173	61	30	23	18	11	2	1076
	0.12	0.13	0.13	0.13	0.07	0.09	0.22	0.36	0.00	0.13
	1.83	1.82	2.16	2.77	3.00	2.52	1.94	1.55	2.00	1.98
2017	368	317	140	57	35	15	26	15	6	979
	0.10	0.09	0.06	0.07	0.09	0.00	0.12	0.00	0.00	0.08
	0.91	0.93	1.36	2.30	2.14	2.53	0.92	1.00	4.00	1.15
2018	343	268	106	38	17	15	16	9	2	814

	Age Group at enlistment (AG0)									
Year of Entry	**16-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**Total**
	0.03 0.00	0.04 0.01	0.08 0.43	0.03 1.61	0.06 0.00	0.07 0.87	0.00 0.00	0.00 1.44	0.00 0.00	0.04 0.17
Total	9,759 0.54 5.32	6,741 0.49 5.24	2,458 0.44 5.34	1,113 0.47 5.74	634 0.47 5.95	588 0.40 6.21	361 0.46 5.40	142 0.43 4.42	24 0.50 4.08	21,820 0.50 5.36

Note: The top number in each cell represents the frequency or number of observations, the second number is the mean failure rate as at 31 May 2018, and the third number is the average years of service.

Table 31 shows the statistics related to the year of entry cohort (each row) and the age group at enlistment (each column). Each cell represents an intersection between a year of entry cohort and one of the age groups. To interpret one of these results, the top left hand corner cell represents all individuals who enlisted into the RAN in 2002 and was in the age group 16–19 at the time. There were 265 individuals in this group, and between their enlistment date and the end of the dataset (31 May 2018) 85% of these individuals have 'failed' (i.e., separated from the RAN) or conversely 15% continue to serve. These individuals had an average years of service of 6.82 years.

In addition to this, the row at the bottom of Table 31 displays the totals for each column. The first total cell contains the total number of observations in age group 16–19 (9,759), the mean failure rate of 0.54 (separation rate of 54%) and the average years of service of 5.32 years. As can be seen by comparing the failure rates across this bottom row, the younger age groups have higher separation rates than the older age groups (except for 55–59 years old).

Noting these results, Tables 30 and 31 indicate the significance of age group at enlistment and cohort year of enlistment in accounting for separation behavior of RAN personnel.

6. Final—Preferred Model

a. All Significant Characteristic Variables

In this section I consolidate all of the characteristic variables which appear to account for separation behavior of RAN personnel. These include: the aggregated function at enlistment (RevisedFunction0), aggregated function at separation (RevisedFunction1), Sailor rank at enlistment (SRank0), Sailor rank at separation (SRank1), Officer rank at separation (ORank1), gender (i.female*i.cumLOS), cohort year of enlistment (year) and age group at enlistment (AG0). Unfortunately, including all these variables simultaneously would result in multicollinearity and the model couldn't be estimated (e.g. 'no observations' error).

b. Most Significant Characteristics Variables

Due to the model specification error above, I then estimated a model which includes all of the most significant variables. These reduced set of variables includes: the aggregated function at enlistment (RevisedFunction0), aggregated function at separation (RevisedFunction1), Sailor rank at enlistment (SRank0), gender (i.female*i.cumLOS), cohort year of enlistment (year) and age group at enlistment (AG0).

Table 32 shows the output for this reduced model specification.

Table 32. Regression output—Most significant characteristics variables.

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
Aggregated Function at Enlistment (RevisedFunction0)	Aeronautical Engineer	_IRevisedFu_2	1.000 (omitted)	
	Air Technician Aircraft	_IRevisedFu_3	4.599 (1.337)	***
	Air Technician Avionics	_IRevisedFu_4	4.483 (1.311)	***
	Aircrew	_IRevisedFu_5	6.803 (4.915)	**
	Aviation Support	_IRevisedFu_6	19.988 (9.858)	***
	Boatswains Mate	_IRevisedFu_7	4.114 (1.108)	***
	Chaplain	_IRevisedFu_8	1.000 (omitted)	
	Clearance Diver	_IRevisedFu_9	4.344 (1.295)	***
	Combat Systems Operator	_IRevisedFu_10	3.812 (0.984)	***
	Combat Systems Operator Mine Warfare	_IRevisedFu_11	4.012 (1.242)	***
	Communications Information Systems	_IRevisedFu_12	3.588 (1.030)	***

Reduced Model		**Interaction Variable**	**Hazard Ratio (s.e.)**	
	Communications Information Systems Submariner	_IRevisedFu_13	3.931 (1.747)	**
	Cryptologic Linguist	_IRevisedFu_14	5.223 (1.583)	***
	Cryptologic Systems	_IRevisedFu_15	5.514 (1.690)	***
	Dental	_IRevisedFu_16	4.229 (1.947)	**
	Dentist	_IRevisedFu_17	1.000 (omitted)	
	Electronic Warfare	_IRevisedFu_18	5.390 (1.817)	***
	Electronic Warfare Submarines	_IRevisedFu_19	4.840 (1.941)	***
	Electronics Technician	_IRevisedFu_20	4.088 (1.105)	***
	Electronics Technician Submariner	_IRevisedFu_21	4.510 (1.571)	***
	General Experience	_IRevisedFu_22	3.897 (1.387)	***
	Hydrographic Systems Operator	_IRevisedFu_23	3.478 (1.086)	***
	Imagery Specialist	_IRevisedFu_24	2.849 (3.326)	

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
	Intelligence	_IRevisedFu_25	1.000 (omitted)	
	Legal	_IRevisedFu_26	1.000 (omitted)	
	Management Executive	_IRevisedFu_27	1.000 (omitted)	
	Marine Engineer	_IRevisedFu_28	1.000 (omitted)	
	Marine Engineer Submariner	_IRevisedFu_29	1.000 (omitted)	
	Marine Technician	_IRevisedFu_30	3.841 (1.035)	***
	Marine Technician Submariner	_IRevisedFu_31	5.445 (1.709)	***
	Maritime Aviation Warfare Officer	_IRevisedFu_32	1.000 (omitted)	
	Maritime Geospatial Officer (Hydrographer)	_IRevisedFu_33	1.000 (omitted)	
	Maritime Geospatial Officer (Meteorologist/Oceanographer)	_IRevisedFu_34	1.000 (omitted)	
	Maritime Logistics Chef	_IRevisedFu_35	4.012 (1.110)	***
	Maritime Logistics Chef Submariner	_IRevisedFu_36	3.808 (2.192)	*

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
	Maritime Logistics Officer	_IRevisedFu_37	1.000 (omitted)	
	Maritime Logistics Personnel Operations	_IRevisedFu_38	3.793 (1.141)	***
	Maritime Logistics Supply Chain	_IRevisedFu_39	2.880 (0.846)	***
	Maritime Logistics Supply Chain Submariner	_IRevisedFu_40	0.000 (0.000)	
	Maritime Logistics Support Operations	_IRevisedFu_41	3.521 (1.000)	***
	Maritime Logistics Support Operations Submariner	_IRevisedFu_42	5.297 (4.340)	*
	Maritime Warfare Officer	_IRevisedFu_43	1.000 (omitted)	
	Maritime Warfare Officer Submariner	_IRevisedFu_44	1.000 (omitted)	
	Medical	_IRevisedFu_45	4.372 (1.320)	***
	Medical Administration	_IRevisedFu_46	1.000 (omitted)	
	Medical Officer	_IRevisedFu_47	1.000 (omitted)	
	Medical Submariner	_IRevisedFu_48	3.134 (3.863)	

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
	Mine Clearance Diver	_IRevisedFu_49	1.000 (omitted)	
	Musician	_IRevisedFu_50	0.000 (0.000)	***
	Naval Police Coxswain (Officer)	_IRevisedFu_51	1.000 (omitted)	
	Naval Police Coxswain (Sailor)	_IRevisedFu_52	3.336 (1.176)	***
	Nurse	_IRevisedFu_53	1.000 (omitted)	
	Other Officers	_IRevisedFu_54	1.000 (omitted)	
	Other Sailors	_IRevisedFu_55	4.620 (1.237)	***
	Physical Trainer	_IRevisedFu_56	5.311 (2.920)	**
	Pilot	_IRevisedFu_57	1.000 (omitted)	
	Principal Warfare Officer	_IRevisedFu_58	1.000 (omitted)	
	Training Systems	_IRevisedFu_59	1.000 (omitted)	
	Warrant Officer (Entry)	_IRevisedFu_60	1.000 (omitted)	

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)
	Weapons Electrical Aircraft Engineer	_IRevisedFu_61	1.000 (omitted)
	Weapons Electrical Engineer	_IRevisedFu_62	1.000 (omitted)
	Weapons Electrical Engineer Submariner	_IRevisedFu_63	1.000 (omitted)
Aggregated Function at Separation (RevisedFunction1)	Aeronautical Engineer	_IRevisedFua2	0.000 (0.000)
	Air Technician Aircraft	_IRevisedFua3	1.239 (0.301)
	Air Technician Avionics	_IRevisedFua4	1.121 (0.281)
	Aircrew	_IRevisedFua5	1.038 (0.357)
	Aviation Support	_IRevisedFua6	0.499 (0.195)
	Band	_IRevisedFua7	0.000 (36.199)
	Boatswains Mate	_IRevisedFua8	1.524 (0.330)
	Chaplain	_IRevisedFua9	1.000 (omitted)
	Clearance Diver	_IRevisedFua10	1.103 (0.279)

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)
	Combat Systems Operator	_IRevisedFua11	1.269 (0.260)
	Combat Systems Operator Mine Warfare	_IRevisedFua12	1.220 (0.325)
	Combat Systems Operator-Above	_IRevisedFua13	0.000 (0.000)
	Combat Systems Operator-Above-ASAC	_IRevisedFua14	0.000 (0.000)
	Combat Systems Operator-Under	_IRevisedFua15	0.000 (0.000)
	Communications Information Systems	_IRevisedFua16	1.509 (0.360)
	Communications Information Systems Submariner	_IRevisedFua17	1.310 (0.435)
	Cryptologic Linguist	_IRevisedFua18	1.223 (0.324)
	Cryptologic Systems	_IRevisedFua19	1.262 (0.345)
	Dental	_IRevisedFua20	2.208 (0.917)
	Dentist	_IRevisedFua21	1.000 (omitted)
	Electronic Warfare	_IRevisedFua22	1.287 (0.314)

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
	Electronic Warfare Submarines	_IRevisedFua23	1.344	
			(0.454)	
	Electronics Technician	_IRevisedFua24	1.285	
			(0.282)	
	Electronics Technician Submariner	_IRevisedFua25	1.059	
			(0.303)	
	General Experience	_IRevisedFua26	1.167	
			(0.489)	
	Hydrographic Systems Operator	_IRevisedFua27	1.717	*
			(0.437)	
	Imagery Specialist	_IRevisedFua28	1.386	
			(0.542)	
	Intelligence	_IRevisedFua29	0.256	
			(0.262)	
	Legal	_IRevisedFua30	1.000	
			(omitted)	
	Management Executive	_IRevisedFua31	1.169	
			(1.205)	
	Marine Engineer	_IRevisedFua32	0.660	
			(0.408)	
	Marine Engineer Submariner	_IRevisedFua33	0.000	
			(0.000)	
	Marine Technician	_IRevisedFua34	1.484	
			(0.322)	

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)
	Marine Technician Submariner	_IRevisedFua35	1.544 (0.378)
	Maritime Aviation Warfare Officer	_IRevisedFua36	0.000 (0.000)
	Maritime Geospatial Officer (Hydrographer)	_IRevisedFua37	3.811 (3.967)
	Maritime Geospatial Officer (Meteorologist/Oceanographer)	_IRevisedFua38	1.000 (omitted)
	Maritime Logistics Chef	_IRevisedFua39	1.514 (0.340)
	Maritime Logistics Chef Submariner	_IRevisedFua40	1.589 (0.614)
	Maritime Logistics Officer	_IRevisedFua41	0.813 (0.402)
	Maritime Logistics Personnel Operations	_IRevisedFua42	1.347 (0.333)
	Maritime Logistics Supply Chain	_IRevisedFua43	1.474 (0.356)
	Maritime Logistics Supply Chain Submariner	_IRevisedFua44	0.282 (0.293)
	Maritime Logistics Support Operations	_IRevisedFua45	1.424 (0.330)
	Maritime Logistics Support Operations Submariner	_IRevisedFua46	1.824 (0.816)

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)
	Maritime Warfare Officer	_IRevisedFua47	1.093 (0.383)
	Maritime Warfare Officer Submariner	_IRevisedFua48	0.000 (0.000)
	Medical	_IRevisedFua49	1.246 (0.314)
	Medical Administration	_IRevisedFua50	0.000 (0.000)
	Medical Officer	_IRevisedFua51	1.759 (1.807)
	Medical Submariner	_IRevisedFua52	0.770 (0.563)
	Mine Clearance Diver	_IRevisedFua53	1.000 (omitted)
	Musician	_IRevisedFua54	1.41E+07 (omitted)
	Naval Police Coxswain (Officer)	_IRevisedFua55	5.295 (5.596)
	Naval Police Coxswain (Sailor)	_IRevisedFua56	1.115 (0.281)
	Nurse	_IRevisedFua57	7.325 (7.823)
	Other Officers	_IRevisedFua58	1.000 (omitted)

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
	Other Sailors	_IRevisedFua59	3.531	***
			(0.856)	
	Physical Trainer	_IRevisedFua60	0.975	
			(0.306)	
	Pilot	_IRevisedFua61	0.000	
			(0.000)	
	Principal Warfare Officer	_IRevisedFua62	0.000	
			(0.000)	
	Principal Warfare Officer Amphib	_IRevisedFua63	1.000	
			(omitted)	
	Senior Officer	_IRevisedFua64	1.000	
			(omitted)	
	Training Systems	_IRevisedFua65	0.631	
			(0.644)	
	Warrant Officer (Entry)	_IRevisedFua66	1.675	
			(1.266)	
	Weapons Electrical Aircraft Engineer	_IRevisedFua67	0.000	
			(0.000)	
	Weapons Electrical Engineer	_IRevisedFua68	1.238	
			(0.514)	
	Weapons Electrical Engineer Submariner	_IRevisedFua69	0.000	
			(0.000)	
Sailor Rank at enlistment	E01	_ISRank0_2	0.762	**

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
(SRank0)			(0.072)	
	E02	_ISRank0_3	0.858	
			(0.175)	
	E03	_ISRank0_4	1.178	*
			(0.092)	
	E05	_ISRank0_5	1.431	***
			(0.123)	
	E06	_ISRank0_6	1.103	
			(0.131)	
	E08	_ISRank0_7	0.944	
			(0.147)	
	E09	_ISRank0_8	0.653	
			(0.656)	
Gender interacted with cumulative LOS	female=1,cumLOS=1	_IfemXcum_1_1	0.989	
			(0.098)	
(i.female*i.cumLOS)	female=1,cumLOS=2	_IfemXcum_1_2	1.218	
			(0.138)	
	female=1,cumLOS=3	_IfemXcum_1_3	1.255	*
			(0.139)	
	female=1,cumLOS=4	_IfemXcum_1_4	1.180	
			(0.109)	
	female=1,cumLOS=5	_IfemXcum_1_5	0.940	
			(0.098)	
	female=1,cumLOS=6	_IfemXcum_1_6	0.949	

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
			(0.086)	
	female=1,cumLOS=7	_IfemXcum_1_7	0.998	
			(0.111)	
	female=1,cumLOS=8	_IfemXcum_1_8	0.898	
			(0.105)	
	female=1,cumLOS=9	_IfemXcum_1_9	1.160	
			(0.153)	
	female=1,cumLOS=10	_IfemXcum_1_10	0.691	*
			(0.107)	
	female=1,cumLOS=11	_IfemXcum_1_11	0.946	
			(0.159)	
	female=1,cumLOS=12	_IfemXcum_1_12	0.924	
			(0.189)	
	female=1,cumLOS=13	_IfemXcum_1_13	0.780	
			(0.224)	
	female=1,cumLOS=14	_IfemXcum_1_14	1.041	
			(0.354)	
	female=1,cumLOS=15	_IfemXcum_1_15	1.024	
			(0.511)	
	female=1,cumLOS=16	_IfemXcum_1_16	1.00E+10	
			(omitted)	
Calendar year of enlistment	2002	_Iyear_2002	2.078	***
Cohort			(0.123)	

Reduced Model		**Interaction Variable**	**Hazard Ratio (s.e.)**	
(year)	2003	_Iyear_2003	1.381	***
			(0.065)	
	2004	_Iyear_2004	1.307	***
			(0.063)	
	2005	_Iyear_2005	1.269	***
			(0.067)	
	2006	_Iyear_2006	1.216	***
			(0.062)	
	2007	_Iyear_2007	1.137	**
			(0.056)	
	2009	_Iyear_2009	0.974	
			(0.050)	
	2010	_Iyear_2010	0.668	***
			(0.038)	
	2011	_Iyear_2011	0.755	***
			(0.045)	
	2012	_Iyear_2012	0.485	***
			(0.037)	
	2013	_Iyear_2013	0.410	***
			(0.030)	
	2014	_Iyear_2014	0.215	***
			(0.016)	
	2015	_Iyear_2015	0.087	***
			(0.008)	

Reduced Model		Interaction Variable	Hazard Ratio (s.e.)	
	2016	_Iyear_2016	0.117	***
			(0.013)	
	2017	_Iyear_2017	0.006	***
			(0.001)	
	2018	_Iyear_2018	0.025	***
			(0.005)	
Age Group at Enlistment (AG0)	Age2024	_IAG0_2	0.909	***
			(0.022)	
	Age2529	_IAG0_3	0.660	***
			(0.028)	
	Age3034	_IAG0_4	0.754	***
			(0.044)	
	Age3539	_IAG0_5	0.321	***
			(0.027)	
	Age4044	_IAG0_6	0.866	
			(0.077)	
	Age4549	_IAG0_7	0.969	
			(0.108)	
	Age5054	_IAG0_8	1.101	
			(0.195)	
	Age5559	_IAG0_9	1.846	
			(1.086)	

Reduced Model	Interaction Variable	Hazard Ratio (s.e.)
	N	21820

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline groups are ‘Acoustic Warfare Analyst’ for aggregate function at enlistment (RevisedFunction0) and aggregate function at separation (RevisedFunction1), ‘E00’ for Sailor rank at enlistment (SRank0), ‘Male’ for interacted term: gender & cumulative LOS (i.female*i.cumLOS), ‘2008’ for calendar year of entry cohort (year) and ‘Age1619’ for age group at enlistment (AG0).

Other regressors include i.cumLOS and i.female, however were excluded from the output table for formatting purposes.

When including the additional variables (multivariate) into the model, compared with the single variable (bivariate) predictive models, some of the variables have become less statistically significant. Presumably this is due to the mediating effects of the combinations of other variables. These include aggregated function at separation (RevisedFunction1), which was 26% not statistically significant in the single model versus 97% not statistically significant in the combined model.

Sailor rank at enlistment (SRank0) was 29% not statistically significant in the single model versus 57% not statistically significant in the combined model.

The interaction between the two dummies of female and cumulative length of service (i.female*i.cumLOS) was 50% not statistically significant in the single model versus 88% not statistically significant in the combined model.

Age group at enlistment (AG0) was 13% not statistically significant in the single model versus 50% not statistically significant in the combined model.

Cohort year of enlistment (year) remained the same with 6% not statistically significant for both the single and combined models.

In contrast, aggregated function at enlistment (RevisedFunction0) became more statistically significant, with 26% not statistically significant in the single model versus 5% not statistically significant in the combined (multivariate) model.

Table 33 summarizes the statistical significance changes between the single (bivariate) and combined (multivariate) predictive models for each characteristic variable.

Table 33. Summary of statistically significance changes for all variables between single and combined models.

Variables	Single (Bivariate) model	Combined (Multivariate) model	Change
Aggregated Function at enlistment (RevisedFunction0)	26% not stat sig	5% not stat sig	Improved by 21%
Aggregated Function at separation (RevisedFunction1),	26% not stat sig	97% not stat sig	Worsened by 71%
Sailor rank at enlistment (SRank0)	29% not stat sig	57% not stat sig	Worsened by 28%
Interaction between the two dummies of female and cumulative length of service (i.female*i.cumLOS)	50% not stat sig	88% not stat sig	Worsened by 38%
Cohort year of enlistment (year)	6% not stat sig	6% not stat sig	No change
Age group at enlistment (AG0)	13% not stat sig	50% not stat sig	Worsened by 37%

The fact that aggregated function at enlistment (RevisedFunction0) became more statistically significant in the final model highlights that this variable is an extremely important variable in this analysis, and further detailed research of this variable interacted with other key variables (like gender, age group, and cohort year of enlistment) is highly recommended.

C. SENSITIVITY ANALYSES AND VALIDATION

1. Test Statistics for Cox

a. Revised Model Specification from Previous Section

With a final set of predictors (the full preferred model) in the Cox model, I will now conduct postestimation tests on the model's validity. I first implement a link test which seeks to determine whether the model is correctly specified. The link test's null hypothesis is that the Cox full model specification is 'correct', that is, there are no significant omitted variables. The test is based on re-estimation where some added set of variables will add little to no explanatory power than what is currently included. Thus, I want to see a 'high' p-value on the added variables, meaning we fail to reject the null.

Using the full preferred model, I found a p-value of 0.000 from the link test indicating that the predictive model could be improved upon.

b. Best (and Valid) Model Specification

I then estimated various specifications, and verified that the following model is a valid specification, with a hatsq p-value (0.563), whilst still including the most important variables: the aggregated function at enlistment (RevisedFunction0), aggregated function at separation (RevisedFunction1), Sailor rank at enlistment (SRank0), and gender (i.female*i.cumLOS).

The test statistic on hatsq (p=0.563) suggests that the model specification including these controls is valid. Going forward this will be the model specification which I will use to further validate and compare this Cox model with other analysis methodologies. Table 34 contains the regression output for this best and valid model.

Table 34. Regression output—Best and valid model specification.

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
Aggregated Function at Enlistment (RevisedFunction0)	Aeronautical Engineer	_IRevisedFu_2	1.000 (omitted)	
	Air Technician Aircraft	_IRevisedFu_3	3.893 (1.101)	***
	Air Technician Avionics	_IRevisedFu_4	3.840 (1.080)	***
	Aircrew	_IRevisedFu_5	3.549 (2.507)	
	Aviation Support	_IRevisedFu_6	3.885 (1.836)	**
	Boatswains Mate	_IRevisedFu_7	3.479 (0.887)	***
	Chaplain	_IRevisedFu_8	1.000 (omitted)	
	Clearance Diver	_IRevisedFu_9	3.843 (1.105)	***
	Combat Systems Operator	_IRevisedFu_10	3.032 (0.744)	***

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
	Combat Systems Operator Mine Warfare	_IRevisedFu_11	1.927	
			(0.678)	
	Communications Information Systems	_IRevisedFu_12	2.670	***
			(0.733)	
	Communications Information Systems Submariner	_IRevisedFu_13	1.499	
			(0.650)	
	Cryptologic Linguist	_IRevisedFu_14	3.697	***
			(1.086)	
	Cryptologic Systems	_IRevisedFu_15	3.467	***
			(1.051)	
	Dental	_IRevisedFu_16	3.368	*
			(1.680)	
	Dentist	_IRevisedFu_17	1.000	
			(omitted)	
	Electronic Warfare	_IRevisedFu_18	1.345	
			(0.441)	
	Electronic Warfare Submarines	_IRevisedFu_19	2.400	*
			(0.926)	
	Electronics Technician	_IRevisedFu_20	3.863	***
			(0.994)	
	Electronics Technician Submariner	_IRevisedFu_21	2.399	*
			(0.830)	
	General Experience	_IRevisedFu_22	2.192	*

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
			(0.736)	
	Hydrographic Systems Operator	_IRevisedFu_23	2.059	*
			(0.614)	
	Imagery Specialist	_IRevisedFu_24	2.174	
			(2.397)	
	Intelligence	_IRevisedFu_25	1.000	
			(omitted)	
	Legal	_IRevisedFu_26	1.000	
			(omitted)	
	Management Executive	_IRevisedFu_27	1.000	
			(omitted)	
	Marine Engineer	_IRevisedFu_28	1.000	
			(omitted)	
	Marine Engineer Submariner	_IRevisedFu_29	1.000	
			(omitted)	
	Marine Technician	_IRevisedFu_30	2.908	***
			(0.750)	
	Marine Technician Submariner	_IRevisedFu_31	1.762	
			(0.540)	
	Maritime Aviation Warfare Officer	_IRevisedFu_32	1.000	
			(omitted)	
	Maritime Geospatial Officer (Hydrographer)	_IRevisedFu_33	1.000	
			(omitted)	
	Maritime Geospatial Officer (Meteorologist/Oceanographer)	_IRevisedFu_34	1.000	

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
			(omitted)	
	Maritime Logistics Chef	_IRevisedFu_35	2.924	***
			(0.778)	
	Maritime Logistics Chef Submariner	_IRevisedFu_36	1.236	
			(0.740)	
	Maritime Logistics Officer	_IRevisedFu_37	1.000	
			(omitted)	
	Maritime Logistics Personnel Operations	_IRevisedFu_38	2.439	**
			(0.716)	
	Maritime Logistics Supply Chain	_IRevisedFu_39	1.998	*
			(0.566)	
	Maritime Logistics Supply Chain Submariner	_IRevisedFu_40	0.000	
			(0.000)	
	Maritime Logistics Support Operations	_IRevisedFu_41	2.070	**
			(0.568)	
	Maritime Logistics Support Operations Submariner	_IRevisedFu_42	1.195	
			(0.997)	
	Maritime Warfare Officer	_IRevisedFu_43	1.000	
			(omitted)	
	Maritime Warfare Officer Submariner	_IRevisedFu_44	1.000	
			(omitted)	
	Medical	_IRevisedFu_45	3.249	***
			(0.949)	
	Medical Administration	_IRevisedFu_46	1.000	

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
			(omitted)	
	Medical Officer	_IRevisedFu_47	1.000	
			(omitted)	
	Medical Submariner	_IRevisedFu_48	1.254	
			(1.491)	
	Mine Clearance Diver	_IRevisedFu_49	1.000	
			(omitted)	
	Musician	_IRevisedFu_50	0.000	***
			(0.000)	
	Naval Police Coxswain (Officer)	_IRevisedFu_51	1.000	
			(omitted)	
	Naval Police Coxswain (Sailor)	_IRevisedFu_52	4.487	***
			(1.509)	
	Nurse	_IRevisedFu_53	1.000	
			(omitted)	
	Other Officers	_IRevisedFu_54	1.000	
			(omitted)	
	Other Sailors	_IRevisedFu_55	3.462	***
			(0.881)	
	Physical Trainer	_IRevisedFu_56	6.968	***
			(3.741)	
	Pilot	_IRevisedFu_57	1.000	
			(omitted)	
	Principal Warfare Officer	_IRevisedFu_58	1.000	

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
			(omitted)	
	Training Systems	_IRevisedFu_59	1.000	
			(omitted)	
	Warrant Officer (Entry)	_IRevisedFu_60	1.000	
			(omitted)	
	Weapons Electrical Aircraft Engineer	_IRevisedFu_61	1.000	
			(omitted)	
	Weapons Electrical Engineer	_IRevisedFu_62	1.000	
			(omitted)	
	Weapons Electrical Engineer Submariner	_IRevisedFu_63	1.000	
			(omitted)	
Aggregated Function at Separation (RevisedFunction1)	Aeronautical Engineer	_IRevisedFua2	0.000	
			(0.000)	
	Air Technician Aircraft	_IRevisedFua3	0.760	
			(0.186)	
	Air Technician Avionics	_IRevisedFua4	0.698	
			(0.173)	
	Aircrew	_IRevisedFua5	0.760	
			(0.256)	
	Aviation Support	_IRevisedFua6	0.387	*
			(0.145)	
	Band	_IRevisedFua7	0.000	
			(35.961)	
	Boatswains Mate	_IRevisedFua8	0.994	

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
			(0.209)	
	Chaplain	_IRevisedFua9	1.000	
			(omitted)	
	Clearance Diver	_IRevisedFua10	0.808	
			(0.202)	
	Combat Systems Operator	_IRevisedFua11	1.026	
			(0.206)	
	Combat Systems Operator Mine Warfare	_IRevisedFua12	0.479	*
			(0.155)	
	Combat Systems Operator-Above	_IRevisedFua13	0.000	
			(0.000)	
	Combat Systems Operator-Above-ASAC	_IRevisedFua14	0.000	
			(0.000)	
	Combat Systems Operator-Under	_IRevisedFua15	0.000	
			(0.000)	
	Communications Information Systems	_IRevisedFua16	1.172	
			(0.274)	
	Communications Information Systems Submariner	_IRevisedFua17	0.872	
			(0.286)	
	Cryptologic Linguist	_IRevisedFua18	0.927	
			(0.244)	
	Cryptologic Systems	_IRevisedFua19	0.864	
			(0.242)	
	Dental	_IRevisedFua20	1.438	

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)
			(0.660)
	Dentist	_IRevisedFua21	1.000
			(omitted)
	Electronic Warfare	_IRevisedFua22	0.952
			(0.230)
	Electronic Warfare Submarines	_IRevisedFua23	0.875
			(0.288)
	Electronics Technician	_IRevisedFua24	0.802
			(0.172)
	Electronics Technician Submariner	_IRevisedFua25	0.592
			(0.171)
	General Experience	_IRevisedFua26	1.431
			(0.595)
	Hydrographic Systems Operator	_IRevisedFua27	1.458
			(0.362)
	Imagery Specialist	_IRevisedFua28	1.010
			(0.393)
	Intelligence	_IRevisedFua29	0.178
			(0.182)
	Legal	_IRevisedFua30	1.000
			(omitted)
	Management Executive	_IRevisedFua31	0.806
			(0.830)
	Marine Engineer	_IRevisedFua32	0.398

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)
			(0.245)
	Marine Engineer Submariner	_IRevisedFua33	0.000
			(0.000)
	Marine Technician	_IRevisedFua34	1.067
			(0.228)
	Marine Technician Submariner	_IRevisedFua35	1.099
			(0.268)
	Maritime Aviation Warfare Officer	_IRevisedFua36	0.000
			(0.000)
	Maritime Geospatial Officer (Hydrographer)	_IRevisedFua37	4.460
			(4.629)
	Maritime Geospatial Officer (Meteorologist/Oceanographer)	_IRevisedFua38	1.000
			(omitted)
	Maritime Logistics Chef	_IRevisedFua39	1.284
			(0.287)
	Maritime Logistics Chef Submariner	_IRevisedFua40	0.963
			(0.388)
	Maritime Logistics Officer	_IRevisedFua41	0.608
			(0.298)
	Maritime Logistics Personnel Operations	_IRevisedFua42	1.252
			(0.311)
	Maritime Logistics Supply Chain	_IRevisedFua43	1.078
			(0.258)
	Maritime Logistics Supply Chain Submariner	_IRevisedFua44	0.227

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)
			(0.234)
	Maritime Logistics Support Operations	_IRevisedFua45	1.245
			(0.286)
	Maritime Logistics Support Operations Submariner	_IRevisedFua46	1.427
			(0.652)
	Maritime Warfare Officer	_IRevisedFua47	0.626
			(0.217)
	Maritime Warfare Officer Submariner	_IRevisedFua48	0.000
			(0.000)
	Medical	_IRevisedFua49	0.949
			(0.237)
	Medical Administration	_IRevisedFua50	0.000
			(0.000)
	Medical Officer	_IRevisedFua51	1.187
			(1.218)
	Medical Submariner	_IRevisedFua52	0.467
			(0.331)
	Mine Clearance Diver	_IRevisedFua53	1.000
			(omitted)
	Musician	_IRevisedFua54	2.40E+07
			(omitted)
	Naval Police Coxswain (Officer)	_IRevisedFua55	3.736
			(3.936)
	Naval Police Coxswain (Sailor)	_IRevisedFua56	0.629

Best and Valid Model		**Interaction Variable**	**Hazard Ratio (s.e.)**
			(0.157)
	Nurse	_IRevisedFua57	6.787
			(7.231)
	Other Officers	_IRevisedFua58	1.000
			(omitted)
	Other Sailors	_IRevisedFua59	1.101
			(0.263)
	Physical Trainer	_IRevisedFua60	0.690
			(0.214)
	Pilot	_IRevisedFua61	0.000
			(0.000)
	Principal Warfare Officer	_IRevisedFua62	0.000
			(0.000)
	Principal Warfare Officer Amphib	_IRevisedFua63	1.000
			(omitted)
	Senior Officer	_IRevisedFua64	1.000
			(omitted)
	Training Systems	_IRevisedFua65	0.395
			(0.402)
	Warrant Officer (Entry)	_IRevisedFua66	1.478
			(1.052)
	Weapons Electrical Aircraft Engineer	_IRevisedFua67	0.000
			(0.000)
	Weapons Electrical Engineer	_IRevisedFua68	0.684

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
			(0.282)	
	Weapons Electrical Engineer Submariner	_IRevisedFua69	0.000	
			(0.000)	
Sailor Rank at enlistment (SRank0)	E01	_ISRank0_2	0.714	***
			(0.067)	
	E02	_ISRank0_3	1.080	
			(0.209)	
	E03	_ISRank0_4	0.953	
			(0.073)	
	E05	_ISRank0_5	1.030	
			(0.082)	
	E06	_ISRank0_6	0.843	
			(0.089)	
	E08	_ISRank0_7	0.873	
			(0.117)	
	E09	_ISRank0_8	0.377	
			(0.377)	
Gender interacted with cumulative LOS (i.female*i.cumLOS)	female=1,cumLOS=1	_IfemXcum_1_1	0.518	***
			(0.055)	
	female=1,cumLOS=2	_IfemXcum_1_2	1.094	
			(0.122)	
	female=1,cumLOS=3	_IfemXcum_1_3	2.198	***
			(0.242)	
	female=1,cumLOS=4	_IfemXcum_1_4	1.630	***

Best and Valid Model		Interaction Variable	Hazard Ratio (s.e.)	
			(0.151)	
	female=1,cumLOS=5	_IfemXcum_1_5	1.408	***
			(0.146)	
	female=1,cumLOS=6	_IfemXcum_1_6	1.122	
			(0.102)	
	female=1,cumLOS=7	_IfemXcum_1_7	1.350	**
			(0.151)	
	female=1,cumLOS=8	_IfemXcum_1_8	1.212	
			(0.141)	
	female=1,cumLOS=9	_IfemXcum_1_9	1.644	***
			(0.217)	
	female=1,cumLOS=10	_IfemXcum_1_10	0.970	
			(0.150)	
	female=1,cumLOS=11	_IfemXcum_1_11	1.337	
			(0.225)	
	female=1,cumLOS=12	_IfemXcum_1_12	1.285	
			(0.263)	
	female=1,cumLOS=13	_IfemXcum_1_13	1.121	
			(0.321)	
	female=1,cumLOS=14	_IfemXcum_1_14	1.495	
			(0.509)	
	female=1,cumLOS=15	_IfemXcum_1_15	1.34E+00	
			(0.667)	
	female=1,cumLOS=16	_IfemXcum_1_16	1.27E+10	

Best and Valid Model	Interaction Variable	Hazard Ratio (s.e.)
		(omitted)
	N	21820

Note: Exponentiated coefficients; Standard errors in parentheses

Stars indicate significance: * p<0.05, ** p<0.01, *** p<0.001

Baseline groups are ‘Acoustic Warfare Analyst’ for aggregate function at enlistment (RevisedFunction0) and aggregate function at separation (RevisedFunction1), ‘E00’ for Sailor rank at enlistment (SRank0), and ‘Male’ for interacted term: gender & cumulative LOS (i.female*i.cumLOS).

Other regressors include i.cumLOS and i.female, however were excluded from the output table for formatting purposes.

2. Cox model versus Logit

I next estimated a logit regression model with the same controls as the preferred and valid Cox model above (not reported here, but available upon request).

I then conducted a side-by-side comparison of the model coefficients (hazard ratio for Cox, odds ratio for logit) to confirm the direction of each factor's effect on separation. Both outputs were odds ratios, so this made them easier to compare. An odds ratio less than 1 indicates something is less likely to occur while an odds ratio greater than 1 means that something is more likely to occur.

For a match to occur between each pair they have to be both less than 1, or both greater than 1. I find that 57% of the estimated hazard and odds ratios of the Cox and logit regressions were a match.

In addition to this coefficient comparison, the statistical significance of each pair was also examined. For a match to occur between each pair they have to both have a p-value of less than 0.05 (meaning statistically significant at the 95% level) or both greater than 0.05 (meaning not statistically significant at the 95% level). This comparison resulted in a 42% match of Cox and logit p-values.

However, completing this side-by-side comparison between the Cox and logit models was found to be less valid, which led me to investigate the reason for such a low matching rate. I determined that each model included a different set of regressors for multicollinearity reasons; for example, some Revised Functions were dropped (omitted) in the Cox that were included in the logit and vice versa. The resulting output was thus not based on the same model specifications due to these differing omitted regressors. This is the reason for the low match rate.

I then compared the Log likelihood output for the Cox and logit models. The Cox Log likelihood was -72,813.568 whereas the logit model's Log likelihood result is -11,991.312. The more positive this number the more valid the model, therefore the logit model is a better model for predicting separation (binary dependent variable is fail/survive) across the entire analysis time with the same included characteristics.

However, because of the binary nature of the dependent variable in logit (only fail/survive), the logit does not take into account different probabilities of survival at each duration (LOS), only over the entire duration.

3. Cox Model versus Probit

Next, I estimated a probit (marginal effects) model and the output was compared to the Cox model output (not reported here, but available upon request). Like above, the comparison of model p-values showed 39% of the Cox and probit p-values matched in terms of statistical significance.

Again, like the Cox and logit models comparison, the probit model has dropped (omitted) a different set of regressors than the Cox model and is therefore not directly comparable.

Like above, I then looked at the Log likelihood and found that the probit (-11,993.356) was better than the Cox model (-72,813.568), however the logit model (-11,991.312) was also marginally better than the probit.

4. Cox model versus Kaplan-Meier Model

a. Kaplan-Meier Survival Function (Gender)

Figure 10 displays the Kaplan-Meier survival function for gender. The female curve is the red curve and the male curve is the blue curve. In the early years the gap between the female and male curves indicates that females are less likely to survive (more likely to separate) than males. The gap is increasing in the early years, meaning females are getting more and more likely to separate over time. However, in the later years females become more like males.

This is consistent with what I found in the Cox models where whatever significant gender differences are in the earlier years, compared to years of service greater than 10.

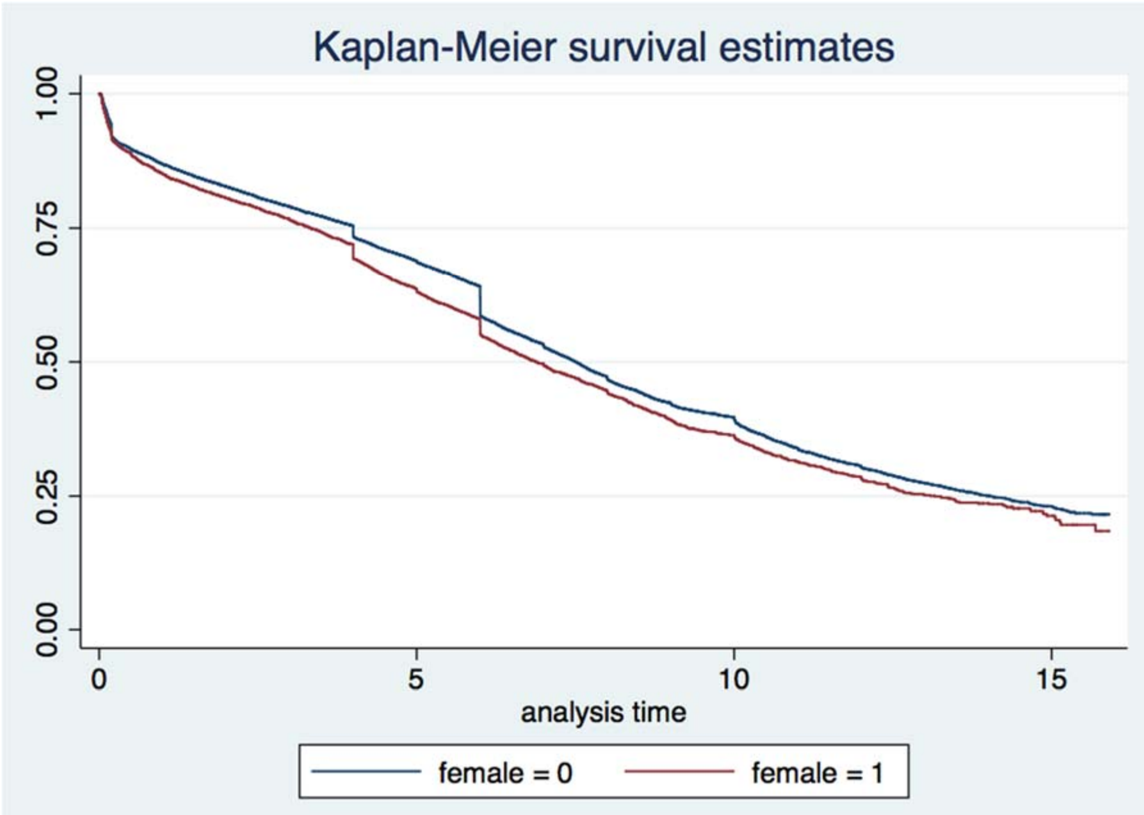


Figure 10. Kaplan-Meier survival function (gender).

b. Kaplan-Meier Hazard Function (Gender)

Figure 11 displays the Kaplan-Meier hazard function by gender. The female curve is the red curve and the male curve is the blue curve. This figure more clearly shows the gap between females and males in the early years and the more similar hazard rates in the later years.

This is consistent with the Cox interacted gender and cumLOS variables model from Table 21. Females more likely to separate than males at earlier years, but not at later years after LOS=10. There is variation occurring at later years at each year, but it appears not significant.

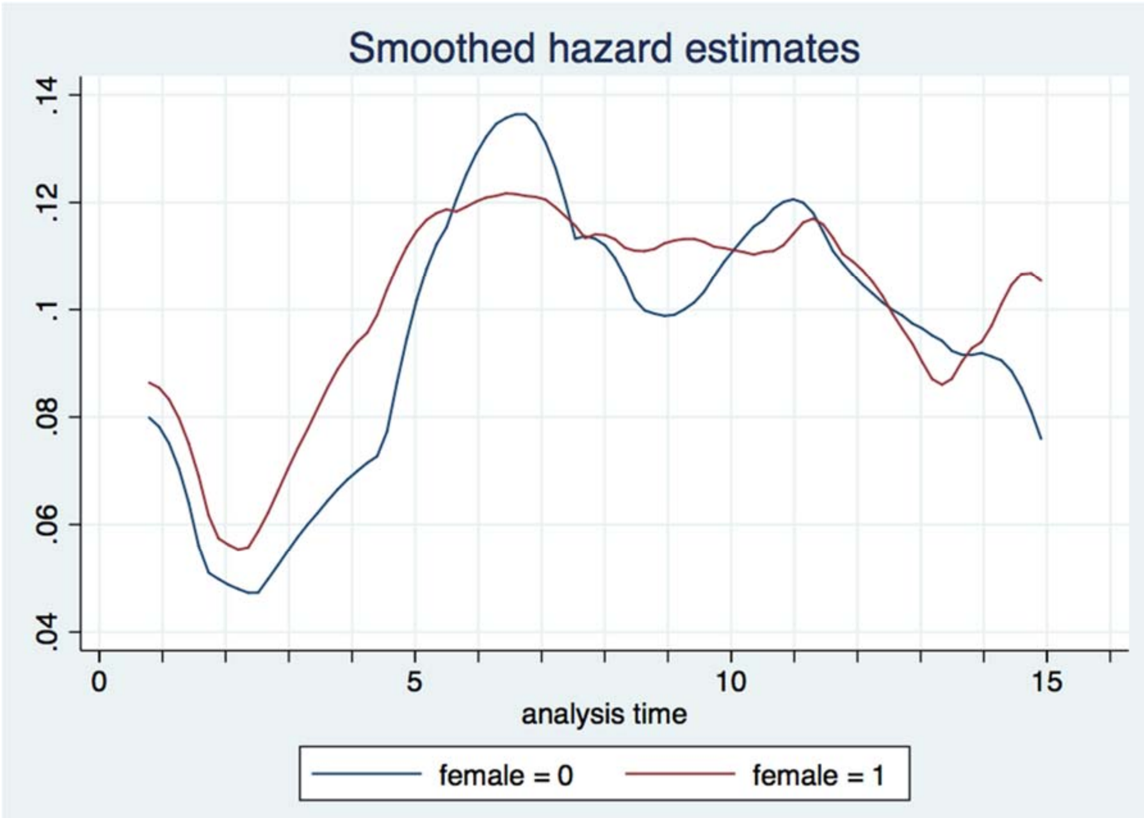


Figure 11. Kaplan-Meier hazard function (gender).

One hypothesis for the highest point on the male curve between LOS six and seven is that this may represent the largest group in the Sailor community, that of 'Marine Technician' and 'Electronics Technician.' They are approximately 93% male and all bound by an IMPS 'contract' of 6 years. All of these Sailors could be completing their first contract at this point and therefore highly likely to separate in the next year.

c. Kaplan-Meier Survival Function (Sailor versus Officer)

Figure 12 displays the Kaplan-Meier survival function for Sailors versus Officers. The Sailors curve is the red curve and the Officers curve is the blue curve. The fact that the Officers curve is above the Sailors curve means that at all time periods Officers are less likely to separate than Sailors. The Sailors curve has distinct drops at LOS=4 and six which are consistent with RAN IMPS 'contract' periods for non-technical and technical Sailors respectively. In addition, the Sailors are significantly more likely to separate than Officers in the first year of service (especially the first six months).

This is consistent with what I found in the raw separation rates in Table 12 where Sailors have a different separation rate profile than Officers and that overall Sailors are more likely to separate at any time in their careers.

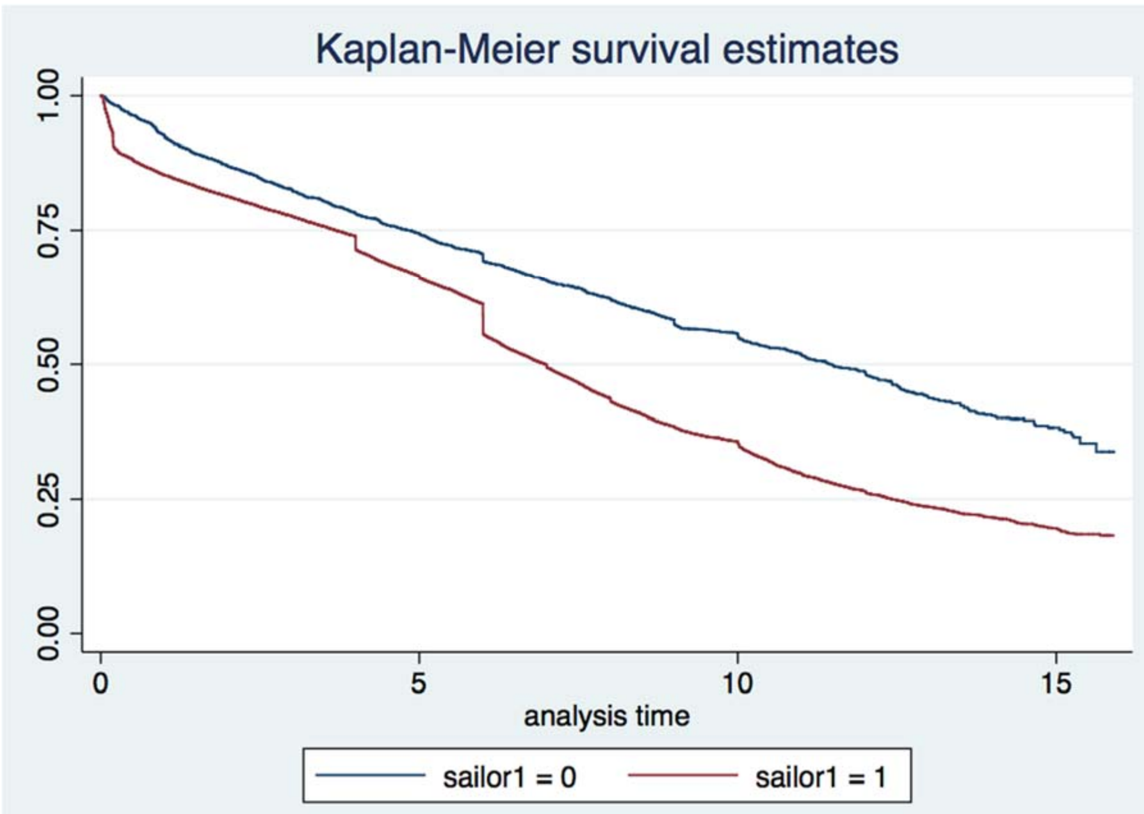


Figure 12. Kaplan-Meier survival function (Sailor versus Officer).

d. *Kaplan-Meier Hazard Function (Sailor versus Officer)*

Figure 13 displays the Kaplan-Meier hazard function for Sailors versus Officers. The Sailors curve is the red curve and the Officers curve is the blue curve. This figure shows that at across all time periods (except LOS 1–3) the instantaneous rate of failure (hazard or separation) for Sailors is greater than for Officers.

Again, this is consistent with Figure 13 and with what I found in the raw separation rates in Table 12.

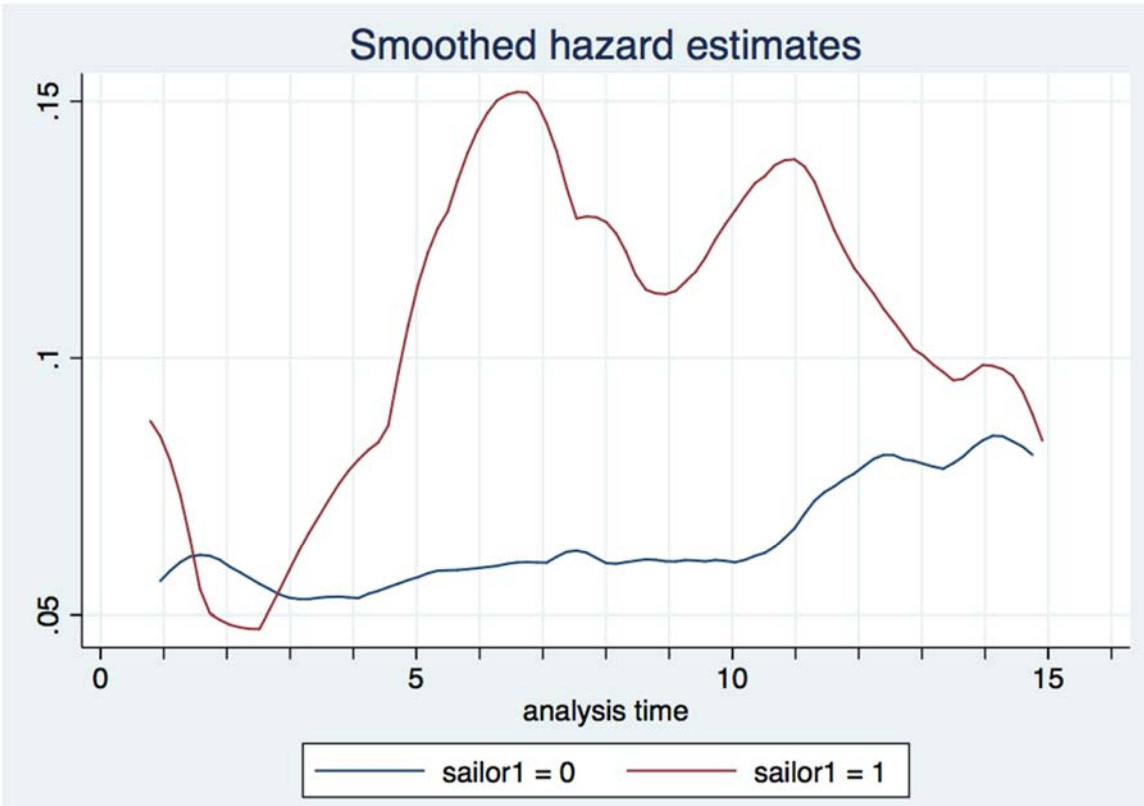


Figure 13. Kaplan-Meier hazard function (Sailor versus Officer).

To determine the more suitable model between Kaplan-Meier and Cox, recall from Chapter II that Kaplan-Meier is suitable at developing survival and hazard functions for models with no covariates or qualitative variables (like gender or Officer versus Sailor). However, Kaplan-Meier is not capable of estimating models with multiple other variables; therefore Cox is better for the research questions this study addresses.

D. CHAPTER SUMMARY

In this chapter I have undertaken a detailed analysis of the dataset and described the most salient findings and results. I conducted an analysis to determine whether using single or multiple spell data was more valid. This was found to be a negligible difference, especially noting the proportion of multiple spell observations in the dataset (approximately 3%). Noting I had multi-spell data I decided to continue using the multi-spell specification going forward. I then developed a survival function curve using Cox for the whole of RAN. The survival profile highlights significant spikes at four and six years of service (length of service) in terms of RAN individuals' separation behavior. I then developed a hazard function curve using Cox for the whole RAN. The hazard profile highlights change in the instantaneous rate of failure/hazard (i.e., separation) at four, six and eight years of service. Next, I conducted an analysis of the separation rates by each cumulative length of service (cumLOS). The first main finding was that the highest separation rate (10.3%) relates to those individuals in the RAN who have not yet completed one year of service. The second finding was that Sailors have a different separation rate pattern/profile than Officers: with high separation rates within the first year of service (11.2%) and at years four and six whereas for Officers the highest is within the first year (5.8%).

The next analysis involved finding the most important characteristics accounting for the survival profiles and the outcome variable, separating (fail) from the service. I conducted a number of different analyses on the key variables of workgroup, rank, gender, age groups, and cohort year of enlistment. These analyses determined that the aggregated function at enlistment (RevisedFunction0), aggregated function at separation (RevisedFunction1), Sailor rank at enlistment (SRank0), interaction between the two dummies of female and cumulative length of service (i.female*i.cumLOS), cohort year of enlistment (i.year), and age group at enlistment (i.AG0) all significantly predict separation behavior for RAN personnel.

Finally, I completed a number of sensitivity analyses and statistical tests to determine whether the model specification is valid. First, I completed a test statistic for the Cox model (link test) on a fully saturated Cox model including all the predictive variables; this model specification was not valid. After further analysis a final valid model was developed and this included aggregated function at enlistment (RevisedFunction0), aggregated function at separation (RevisedFunction1), Sailor rank at enlistment (SRank0), and an interaction between the two dummies of female and cumulative length of service (i.female*i.cumLOS).

Next, I compared the Cox model to a logit regression with the same model specification. This showed that 57% of the coefficients of the Cox and logit regressions were a match and a 42% match of Cox and logit p-values.

In addition, I compared the Cox model to a probit (marginal effects) model with the same specification as above. The result was that 39% of the p-values matched significance between the Cox and probit models.

Having conducted all of the analyses and validations a logit (or probit) model with the preferred and valid specification above is better at predicting separation across the total analysis time. However, the Cox model with the same preferred and valid specification is much better at predicting LOS/cumLOS or duration of service.

At the same time, neither the logit nor probit models would have identified the nuanced relationships between gender, age group or cohort year of enlistment and LOS, as the logit/probit analyze the binary outcome of separation only over the entire LOS. Only a survival or duration analysis model such as the Cox model would have been able to detect varying probabilities of survival at each time period (LOS) among these characteristic dimensions. The analysis here revealed the non-linear effects of gender across individual's LOS. A Cox model determined that there are high probabilities of separation at early LOS for females versus males, and non-statistically significant results in later LOS indicate no significant gender differences in separation behavior after 10 years of service. Both logit and probit models determined an 'average' non statistically significant result of gender in predicting LOS/cumLOS.

Finally, I compared the Cox model outputs to the Kaplan-Meier model outputs for gender and Officer versus Sailor variables. To determine the more suitable model between Kaplan-Meier and Cox, recall from Chapter 2 that Kaplan-Meier is suitable at developing survival and hazard functions for models with no covariates or qualitative variables (like gender or Officer versus Sailor). However, Kaplan-Meier is not capable of estimating models with multiple other variables therefore Cox is better for this use. Figure 14 displays the Kaplan-Meier hazard function by gender side-by-side with the Cox hazard function by gender. In the left plot female is the blue curve and male the red, whereas on the plot on the right female is the red curve and male is the blue curve. Comparing the hazard functions for male indicates that they are practically identical, however the hazard functions for female are similar between cumLOS=0 and four but after that they are very different.

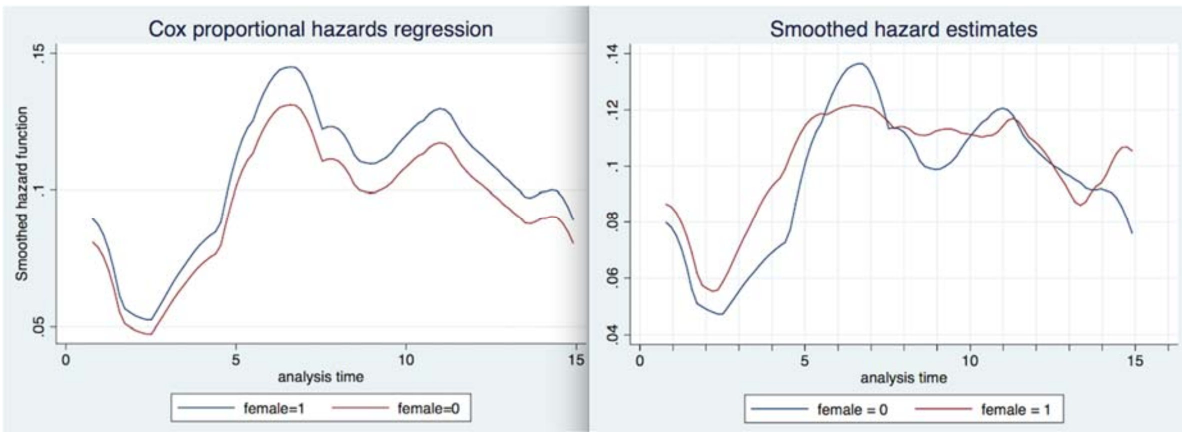


Figure 14. Kaplan-Meier hazard function curve (by gender) versus Cox hazard function curve (by gender).

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION AND RECOMMENDATIONS

A. SUMMARY OF KEY FINDINGS

These are the key findings.

This thesis develops a methodology for preparing survival and hazard (function) profiles for the RAN. These profiles can be implemented along a number of individual characteristic dimensions, including: gender, rank, age group, cohort year of entry and workgroup, or any combination thereof.

The separation rates do vary across cumulative length of service (in years). I found that across the whole RAN the highest likelihood of separation occurs amongst those who have not yet completed one year of service, with a separation rate of 10.3%, followed by six years of service (6.6%) and then four years of service (4.9%). These latter time periods align perfectly with the RAN Initial Minimum Period of Service (IMPS) 'contract' periods, while the early separation indicates potentially bad matches for new RAN members just learning about life in the Navy.

However, I also found that Officers and Sailors have very different separation rates over time. The Sailor's highest three separation rates in order were at zero (11.2%), six (7.2%), and four (5.2%) LOS. The Officer's highest three separation rates in order were at zero (5.8%), one (5.4%) and two (4.1%) LOS. I also found that Officers are more likely to serve for longer periods than Sailors.

Each of the individual characteristics (e.g., gender, age group, rank, cohort year of enlistment, and workgroup) was investigated as predictors, and all were found to significantly account for separation behavior of RAN personnel.

The aggregated workgroups (RevisedFunction) at enlistment and separation were both found to account for separation behavior of RAN personnel. The traditional lower level workgroup (i.e., 'MLO') function (Function) was too granular and needed to be consolidated to include all related workgroup entries (i.e., the RevisedFunction of 'Maritime Logistics Officers' consolidated the Functions of 'MLO,' 'MLO-T' and 'MLO-

UT'). This added power to the analysis because the sample size in each Revised Function was then increased.

I determined that rank does account for separation behavior of RAN personnel, especially Sailor rank at enlistment and separation, in addition to Officer rank at separation.

Both gender and age group were found to have non-linear relationships with length of service (LOS), meaning that the relationship effect varies over the analysis time. Yes, females on average are 10% more likely to separate than males across the whole analysis time; however, and more importantly in the early years of service (2–9 LOS) females are (14–82%) more likely to separate than males, whereas in later years of service females are statistically no different from males with the same years of service. That gender matters more in the earlier years could also coincide with gender differences across the life cycle (i.e., early years are prime childbearing and child rearing years).

Analysis of the calendar year of enlistment was conducted to determine whether cohort and/or peer effects are relevant to the separation behavior of RAN personnel. Two effects were found to be at play: the cohort effect and an economic conditions effect. The cohort effect saw the hazard ratios decline (meaning separation is less likely for later cohorts) as each new year ticked over. The economic conditions effect found that in comparison to 2008 (the baseline year and the year of the Global Financial Crisis (GFC) in Australia), the cohort years 2002–2007 indicate that personnel are more likely to separate than the baseline year, and for cohort years 2009–2018 indicate personnel are less likely to separate. Therefore, the cohort year of enlistment may be a good proxy for macroeconomic conditions.

I found that the youngest age group (16–19 years old) were such outliers in separation behavior compared to all other age groups primarily because they were the largest enlistment age group and had the highest separation rate.

This thesis has also highlighted the strength of a Cox model over either logit or probit models. The Cox model was able to uncover the nuances of the non-linear relationship between gender and length of service, which a logit or probit could not. Under constant model specification characteristics I found that the logit and probit models were better at predicting failure (separation), however, the Cox model was found to be much better at predicting the expected length of service an individual would serve.

I found that the Kaplan-Meier model was perfectly suitable at being used to develop survival or hazard function curves for single qualitative variables only, however the Cox model was essential in developing survival or hazard function profiles when multiple variables are used. For these profiles to be produced using a Cox model, the assumption of conditional proportionality is still required. If the survival or hazard rate is not proportional across multiple groups then Cox may not be the best methodology. This may be the case for Officers versus Sailors and requires further investigation. For future analyses, a Kaplan-Meier survival/hazard curve should be completed for each single variable first (to determine proportionality and validity); and then when combining multiple proportional variables a Cox survival/hazard curve can be completed.

B. CONCLUSIONS AND POLICY CONSIDERATIONS

In light of these findings, the following are my conclusions and recommendations.

If the RAN were to craft policies for retention there are efficiency gains to be had in targeting retention measures at specific years of service. In an environment of constrained resources, if the RAN could more efficiently utilize these limited resources, retention policies could be better targeted towards retention at LOS four, six, as well as gender retention policies in the earlier years.

Gender is an important workforce factor. Should the RAN seek to create a better balance of gender in its workforce, then policy changes targeted to gender diversity need to be focused on women with less than 10 years of service, as women's separation behavior beyond 10 years is not significantly different to men.

Rank is also an important workforce factor. Greater efforts could be expended to incentivize Sailors to choose to remain in the service at the end of their IMPS 'contract' periods of four and six years, rather than train new recruits. In addition, further endeavors could target incentives for Officers to remain in the RAN in their first three years of service, coinciding with their highest rates of separation. Current policy is associated with the highest separation rate in the first year of service for both Officers and Sailors. The RAN should review whether this outcome is desirable given a higher early turnover rate also likely causes increased recruiting costs to the workforce, balanced against the other outcomes such policy achieves (i.e., removing bad workforce matches).

C. FURTHER RESEARCH RECOMMENDATIONS

For further research, I would recommend detailed analysis of the relationship between RevisedFunction0 (and to a lesser extent RevisedFunction1) interacted with other key variables in this analysis (gender, rank, age, and cohort year of enlistment). This would allow the analyst to see, for example, whether the gender differences in separation behavior in early years hold only for certain workgroups or across RAN as a whole. The significant gender differences in separation behavior in the early years of service contrasted to no difference in later years is particularly interesting. Further study could flesh out a more complete picture and aid in formulating gender-specific manpower policies.

I also recommend a further investigation of Officer versus Sailor survival curves using a non-proportional methodology, for example, Kaplan-Meier models or a Machine learning Random Forest algorithm methodology. The Random Forest algorithm is more computationally demanding; however, it may be able to get at these non-proportional differences more accurately and is worthy of investigation.

I also recommend a further investigation be undertaken into other pre-enlistment individual quality characteristics as an avenue to resolve the identified high first year separation rate in addition to the youngest age group effects on separation behavior. Doing this will enable the RAN to determine whether the early turnover rate is acceptable and whether there are other pre-enlistment characteristics which may better predict early separation.

Finally, I recommend that an investigation is conducted to determine whether using the outputs of a Cox model provides significantly better forecast results compared to another form of predicting separation rates. Utilizing the outputs from a survival analysis and bringing that into a separation simulation model (i.e., a Markov chain model) may pay large dividends over the current models used.

LIST OF REFERENCES

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2018). *Statistics for business & economics* (13e revised [edition]). Boston, MA: Cengage learning.

Asch, B. J., Hosek, J. R., & Warner, J. T. (2007). New economics of manpower in the Post–Cold War era. In T. Sandler & K. Hartley (Eds.), *Handbook of defense economics* (Vol. 2, pp. 1075–1138). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S1574-0013(06)02032-1

Asch, B. J., & Warner, J. T. (2001). A Theory of compensation and personnel policy in hierarchical organizations with application to the United States Military. *Journal of labor economics*, *19*(3), 523–562. https://doi.org/10.1086/322072

Australian National Audit Office. (2014). *Recruitment and retention of specialist skills for Navy* (No. 17 2014–15). Retrieved from https://www.anao.gov.au/work/performance-audit/recruitment-and-retention-specialist-skills-navy

Cleves, M. A., Gould, W. W., & Gutierrez, R. G. (2004). *An Introduction to survival analysis using Stata, Revised edition*. Oxfordshire, England: Taylor & Francis. Retrieved from https://books.google.com/books?id=KsJEP-eX-PsC

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187–220.

Gotz, G. A., & McCall, J. J. (1984). *A Dynamic retention model for Air Force Officers: Theory and estimates* (No. RAND/R-3028-AF). Santa Monica, CA: RAND. Retrieved from http://www.dtic.mil/docs/citations/ADA149736

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53*(282), 457–481. https://doi.org/10.1080/01621459.1958.10501452

Kilburn, M. R., & Klerman, J. A. (1999). *Enlistment decisions in the 1990s. Evidence from individual-level data* (No. RAND/MR-944-OSD/A). Santa Monica, CA: RAND. Retrieved from http://www.dtic.mil/docs/citations/ADA373705

Liu, X. (2012). *Survival analysis: Models and applications*. Chichester, West Sussex : [Peking]: Wiley.

Rodney, D. (2017). *Navy manpower planning*. Arlington, VA: CNA.

Rosen, S. (1986). The theory of equalizing differences. In O.C. Ashenfelter & R. Layard (Eds.), *Handbook of labor economics* (Vol. 1, pp. 641–692). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S1573-4463(86)01015-5

Royal Australian Navy. (n.d.-a). About the Royal Australian Navy. Retrieved August 28, 2018, from http://www.navy.gov.au/about-royal-australian-navy

Royal Australian Navy. (n.d.-b). Navy Strategic Command. Retrieved August 28, 2018, from http://www.navy.gov.au/about/organisation/navstratcom

Royal Australian Navy. (n.d.-c). *Navy workforce requirements planning manual* (ANP2101). Canberra: Author.

Royal Australian Navy. (n.d.-d). Australian Defence Force badges of rank and special insignia. Retrieved August 28, 2018, from http://www.navy.gov.au/sites/default/files/documents/ADF_badges_of_rank.pdf

The World Bank. (n.d.). Armed forces personnel, total | Data. Retrieved September 8, 2018, from https://data.worldbank.org/indicator/ms.mil.totl.p1?end=2016&name_desc=false&start=1985&view=chart

Warner, J. T., & Asch, B. J. (1995). The economics of military manpower. In T. Sandler & K. Hartley (Eds.), *Handbook of defense economics* (Vol. 1, pp. 347–398). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S1574-0013(05)80015-8

Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach* (Sixth edition). Boston, MA: Cengage Learning.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California