NPS-GSBPP-19-003

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

RETENTION ANALYSIS MODEL (RAM) FOR NAVY MANPOWER ANALYSIS

by

Jeremy Arkes
Tom Ahn
Amilcar Menichini
William Gates

May 2019

Prepared for: OPNAV N13 Military Personnel Plans and Policy

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE ***(DD-MM-YYYY)***
May 2019

2. REPORT TYPE
Final Technical Report

3. DATES COVERED ***(From-To)***
OCT 2016 to MAR 2018

4. TITLE AND SUBTITLE
RETENTION ANALYSIS MODEL (RAM) FOR NAVY MANPOWER ANALYSIS

5a. CONTRACT NUMBER

5b. GRANT NUMBER
Naval Research Program

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER
NPS-17-N358-A

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Jeremy Arkes
Tom Ahn
Amilcar Menichini
William Gates

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Naval Postgraduate School
School: Graduate School of Business and Public Policy/ Dept: Manpower & Economics
1 University Circle, Monterey CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-GSBPP-19-003

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES
The views expressed in this technical report are those of the authors and do not reflect the official policy or position of the Department of Navy, Department of Defense, or the U.S. Government.

14. ABSTRACT

In the first year of our Retention Analysis Modeling project, we began developing the modelling approach by performing the following analyses: Describe the retention models used to analyze policy levers affecting reenlistment rates, including the simple reenlistment model, the Average Cost of Leaving (ACOL) model, and the Dynamic Retention Model (DRM); critically assess the advantages and disadvantages of these models; indicate what would be needed to obtain more credible estimates from future models; discuss the problems with a one-size-fits-all and one-moment-in-YOS-fits-all mid-career bonuses; Make recommendations for tools to address the questions above, in both setting bonuses in real time and predicting expected impacts in the longer term.

One of our key findings is that existing models suffer from potentially large biases affecting the estimates. The sources of the biases include: reverse causality because a lower reenlistment propensity would lead to higher bonuses; measurement error in correctly coding the bonus considered by the sailor at the time a decision was made; and excess supply because, sometimes, more sailors want to reenlist than are allowed to reenlist.

To minimize biases and maximize their efficacy, we will use the increased computing capacity of modern high-performance computing (HPC) cluster servers, and models will be designed from the ground-up to leverage this increased power; we will incorporate non-monetary incentives, other personalized incentives, and measures of service member quality; we will incorporate new and more detailed data of individual's socio-economic and professional status in the estimation.

15. SUBJECT TERMS
Retention Modeling, Force Structure Modeling; Retention Bonus; Reenlistment Bonus; Logistic Regression; Retention Auction

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
Unclassified	Unclassified	Unclassified	UU	81	William R. Gates
					19b. TELEPHONE NUMBER *(include area code)* 831-656-3621

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Ann E. Rondeau
President

Steven R. Lerman
Provost

The report entitled Retention Analysis Model (RAM) For Navy Manpower Analysis was prepared for OPNAV N13, Military Personnel Plans and Policy and funded by the Naval Research Program at the Naval Postgraduate School.

Further distribution of all or part of this report is authorized

This report was prepared by:

Jeremy Arkes
Associate Professor

Tom Ahn
Assistant Professor

Amilcar Menichini
Associate Professor

William Gates
Professor

Reviewed by:

Keith Snider, Dean
Graduate School of Business and Public Policy

Released by:

Jeffrey D. Paduan
Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

NRP FY18 TECHNICAL REPORT

Retention Analysis Model (RAM) For Navy Manpower Analysis

Primary Investigator:
William R. Gates
Professor of Economics
Graduate School of Business & Public Policy
Naval Postgraduate School

Co-PI:
Jeremy Arkes
Associate Professor of Economics
Graduate School of Business & Public Policy
Naval Postgraduate School

Co-PI:
Tom Ahn
Assistant Professor of Economics
Graduate School of Business & Public Policy
Naval Postgraduate School

Co-PI:
Amilcar Menichini
Associate Professor of Finance
Graduate School of Business & Public Policy
Naval Postgraduate School

ABSTRACT

In the first year of our Retention Analysis Modeling project, we began developing the modelling approach by performing the following analyses: Describe the retention models used to analyze policy levers affecting reenlistment rates, including the simple reenlistment model, the Average Cost of Leaving (ACOL) model, and the Dynamic Retention Model (DRM); critically assess the advantages and disadvantages of these models; indicate what would be needed to obtain more credible estimates from future models; discuss the problems with a one-size-fits-all and one-moment-in-YOS-fits-all mid-career bonuses; Make recommendations for tools to address the questions above, in both setting bonuses in real time and predicting expected impacts in the longer term.

One of our key findings is that existing models suffer from potentially large biases affecting the estimates. The sources of the biases include: reverse causality because a lower reenlistment propensity would lead to higher bonuses; measurement error in correctly coding the bonus considered by the sailor at the time a decision was made; and excess supply because, sometimes, more sailors want to reenlist than are allowed to reenlist.

To minimize biases and maximize their efficacy, we will use the increased computing capacity of modern high-performance computing (HPC) cluster servers, and models will be designed from the ground-up to leverage this increased power; we will incorporate non-monetary incentives, other personalized incentives, and measures of service member quality.

Keywords: *Retention Modeling, Force Structure Modeling; Retention Bonus; Reenlistment Bonus; Logistic Regression; Retention Auction*

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

As the U.S. military evolves to the new Blended Retirement System (BRS), there are several important questions that the Navy needs to determine:

- What is the predicted impact of the BRS on near-term and long-term reenlistment rates?
- How can the Navy tailor retention incentives to individual sub-groups or service members?
- How can the Navy account for and utilize personalized non-monetary incentives in addition to monetary incentives?
- How can the Navy use retention policies to optimize both the quantity and quality of the retained force?
- How should the Navy optimally set the mid-career bonuses to offset any retention effects of the reduced retirement annuity, including amount and timing by sub-group or individual service members?

The first year of our Retention Analysis Modeling project is designed to determine the best steps forward for answering these questions. In particular, we performed the following analyses:

- Describe the various retention models that have been used to analyze policy levers affecting reenlistment rates. These include the simple reenlistment model, the Average Cost of Leaving (ACOL) model, and the Dynamic Retention Model (DRM).
- Critically assess the advantages and disadvantages of these retention models.
- Indicate what would be needed to obtain more credible estimates from future models, particularly dynamic-programming models.
- Discuss the problems with a one-size-fits-all and one-moment-in-YOS-fits-all mid-career bonuses.
- Make recommendations for tools to address the questions above, in both setting bonuses in real time and predicting expected impacts in the longer term.

One of our key findings is that existing models that estimate the effect of policy levers, such as bonuses or retirement-scheme changes, suffer from potentially large biases affecting the estimates. The sources of the biases include the following:

- Reverse causality because a lower reenlistment propensity would lead to higher bonuses
- Measurement error in correctly coding the bonus considered by the sailor at the time a decision was made

- Excess supply because, sometimes, more sailors want to reenlist than are allowed to reenlist

Prior reenlistment models, particularly the DRM and ACOL models, had further issues, in addition to the generic problems just mentioned, including the following:

- High sensitivity of the results to assumptions on discount rates and other matters
- Significant omitted-variables bias from pay differences across years capturing the effects of other factors specific to the time period (e.g., the September 11th attacks and later negative developments with the Iraq War). This means that the pay effects will unwittingly capture the effects of these period-specific effects.

The original reenlistment models, particularly the DRM, were limited by computing power and other matters and were not developed to answer the questions being posed today.

We overview how we plan to push the dynamic programming models forward, to minimize biases and maximize their efficacy:

- We will utilize the increased computing capacity of modern high-performance computing (HPC) cluster servers, and models will be designed from the ground-up to leverage this increased power.
- Non-monetary incentives, other personalized incentives, and measures of service member quality will be incorporated into the model.
- New and more detailed data of individual's socio-economic and professional status will be collected and used in the estimation.

Still, predictive models will be imperfect, and we believe that there are other methods to determine the optimal bonus at a particular point in time. We introduce those other methods in this report. They include the following:

- Designing market auctions to elicit truthful responses to the lowest bonus needed to reenlist for each Sailor
- Designing surveys of people who have already made decisions and have no stake in the game for how they would respond if the incentives (e.g., SRBs) were stronger or weaker
- Aggregate opinions of subject-matter experts, like what was used to find the USS *Scorpion*, the skipjack-class nuclear submarine that sunk in 1968

Based on what we have learned regarding the current modeling efforts, we will continue this research next year by formalizing the modeling approaches that we find will best support the decisions facing military personnel policy-makers, focusing on data-based analysis to support questions regarding military compensation, including the continuation bonus and special pays and retention incentives. We will support the model development each year by focusing on specific questions the model will need to address, for example, considering selected communities identified in conjunction with the project sponsor (e.g., aviation continuation bonuses) or decision elements to be included in the model (e.g., service member quality or non-monetary incentives).

We aim to develop a modeling approach that will improve on and complement prior efforts to model retention. With alternative and more accurate predictions, the model will be able to estimate the retention effects of the new retirement system and any changes in military compensation policy and bonuses.

At the same time, we will continue to explore more market-based approaches to retention and assignment policies. These policies will draw on innovative best practices found in civilian-sector personnel management.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A new military retirement system called the Blended Retirement System (BRS) was implemented in Fiscal Year (FY) 2017, which changed the incentive structure for reenlistments and officer retention, significantly reducing the incentive to stay through 20 years of service. The Department of Defense (DoD) is proposing a bonus at the eight- to 12-year point to compensate for the reduced retention incentive. However, it is unclear which bonus levels would elicit the required retention rates, and it is not clear whether the bonus would vary across time and specialty area. Furthermore, the Navy needs to determine how to retain high-quality sailors with the new incentives, and factor in both monetary and non-monetary compensation incentives.

A "dynamic programming" model developed in the 1980s, the Dynamic Retention Model (DRM), has been the primary tool used to evaluate the impact of proposed talent management policy changes on retention, including the size/timing of continuation bonuses in support of the new BRS. The model produces eight aggregate bonus estimates to achieve desired retention goals for officers and enlisted service members for each of the four services. While the DRM has proven to be useful for making data-driven decisions in setting retention policies, it has limitations in addressing many of the questions being posed by current policy-makers. The DRM only accounts for monetary compensation and does not address non-monetary incentives. In addition, there is no accounting for the quality of sailors recruited or retained.

The DRM and other retention analysis models have important operational limitations, leaving gaps in the Navy's ability to optimally set bonuses. First, they have the crucial empirical problems of any study on the effects of Selective Retention Bonuses (SRBs) on retention: reverse causality (is it bonuses affecting retention and/or retention affecting bonuses?), severe measurement error in identifying the relevant bonus at the true decision point for the service member, and the influence of excess supply due to demand constraints. Regarding the latter, researchers only observe the intersection of demand and supply of service members willing to reenlist, not just the supply, causing data to hide the full effects of various bonuses/pays. Analyses that do not account for these potential issues will generate estimates or predictions that are systematically biased.

This research explores whether prior studies evaluating the effects of bonuses were subject to these biases. It also begins to lay the groundwork for designing a set of approaches to estimate the optimal bonuses across specialty areas and as economic conditions change. These approaches will eventually address the issue of how to keep more talented service members and adjust monetary bonuses for non-monetary compensation, both of which are becoming increasingly important under the current talent management initiatives.

The goal is to develop a research agenda to identify a decision support tool, or suite of tools, that can help the Navy establish the optimal policies and bonuses to meet its retention and quality (talent) requirements. The modeling effort will include both model evaluations and investigations into questions related to retention, the new BRS, officer and enlisted quality, and other talent management initiatives. The answers to the specific questions will support Office of the Chief of Naval Personnel (OPNAV N1) decision-makers in the short run while informing the model requirements in the longer run.

In this initial report, we first describe the various retention models that have been used to estimate the effects of new compensation schemes. We then critically assess these retention models, with a comprehensive accounting of the limitation of these models and the directional biases they are subject to. As the field has advanced in terms of computing power and recognizing sources of bias, we are now more aware of the potential problems and of the role of uncertainty and randomness. In this report, we explore how we can develop a new dynamic programming model to estimate the long-run effects of compensation and retirement policy changes that (a) accounts for uncertainty, (2) minimizes the bias, and (3) maximizes its usefulness in being able to account for quality and non-monetary incentives.

Given the unavoidable limitations and biases of current retention models (including the Dynamic Retention Model), we offer new approaches to set optimal mid-career bonuses and other retention incentives at a particular point in time—optimal in that they are cost effective and are designed to keep the desired quality of talent. The most promising of these new approaches is implementing auction models. We also discuss the actions and decisions that are needed to implement auction models.

II. BACKGROUND

Under the legacy military retirement system, service members received life-long retirement payments provided they spent at least 20 years in service. The amount of this "defined benefit" is a percentage of the average of the highest-36-months of basic pay, with the percentage being 2.5% times the number of years they have served. This would be a 50% multiple for those serving 20 years.

The new Blended Retirement System (BRS) has a combination of a reduced Defined Benefit (from the legacy system) and a Thrift Savings Plan (TSP), which would vest after two years in service. Once in service for 60 days, the TSP will include

- an automatic 1% contribution for the government,
- an automatic 3% contribution for the service member's basic pay, and
- a matching contribution of up to 4% by the government for additional contributions from basic pay.

The present concern is that the Defined Benefit, still vesting at 20 years of service, would only have a multiple on basic pay of 2% times the number of years served. Thus, at 20 years, service members would have an annuity of 40% of their basic pay instead of 50%.

With the legacy system, there were small percentages of service members who would leave the service after 10 or 12 years of service (roughly, a mid-career point), so reenlistment rates are fairly high at that point of their careers. There is less incentive for service members to stay to 20 years of service with the new blended military retirement system because it pays lower annuities after 20 years of service and provides some retirement benefits prior to 20 years of service for service members vested in the TSP. Thus, there should be lower reenlistment rates at the mid-career point.

To address this, the military has implemented a continuation bonus that is given between 8 and 12 years of service, can vary between 2.5 and 15.5 times monthly base pay, and requires at least a three-year service commitment. So far, there has been little justification for how large the bonuses need to be to elicit the desired reenlistment rates. One study by RAND (Asch, Mattock, & Hosek, 2015), based on the Dynamic Retention Model (DRM), was able to produce an optimal bonus by service and enlisted versus officer (thus, eight different bonuses and not distinguished by MOS/rating/skill or community). But, the DRM does not address some of the issues currently of interest: tailoring retention incentives to individual sub-groups or service members, incorporating non-monetary incentives, and considering the quality of the retained force.

III. BASIC DESCRIPTIONS OF RETENTION MODELS

Retention/reenlistment models were designed to examine the effects of various compensation policies on reenlistment rates. These include the following:

- the effects of bonuses
- military pay elasticities (which indicate how increases in pay affect reenlistment rates)
- the effects of retirement-pay changes or other changes to compensation schemes

There have been three basic models of reenlistment that could speak to the issue of how to set the most-efficient bonuses. They are

- the basic retention model
- the Annualized Cost of Leaving (ACOL) model
- the Dynamic Retention Model (DRM)

A. BASIC RETENTION MODELS

These models have mostly been used to estimate the effects of bonuses. They typically involve including a measure of the bonus (e.g., the SRB multiple or SRB bonus amount for enlisted personnel) directly in the model, as in Equation 1:

$$R_{ist} = X_{ist}\beta + \gamma * SRB_{ist} + \varepsilon_{ist} \qquad (1)$$

where

- R_{ist} is the reenlistment decision for sailor *i* with skill *s* in period *t*.
- X is a set of control variables, such as skill (occupation) fixed effects and demographic variables.
- SRB is the multiple or the bonus amount.

This method was used in many of the early studies of military retention (e.g., Hosek & Peterson, 1985). Some earlier studies (e.g., Enns, 1975) were based on occupation averages, so the dependent variable was the reenlistment rate for the occupation in the given period. Arkes (2016) used this in a recent application, with a wide set of fixed effects to attempt to control for the reverse causality that is discussed Chapter IV Sections B and C.

We should note that, in some cases, the basic retention model, with its more agnostic approach toward the preferences and motivations of individuals, may yield more accurate predictions. More complicated models such as ACOL and DRM require explicit, upfront assumptions. If these assumptions are incorrect, estimates derived from these models may be of questionable value. We discuss this further in the following sections.

B. ACOL MODELS

Initially developed by Warner (1979), the Annualized-Cost-of-Leaving (ACOL) model attempts to incorporate nearly all forms of compensation so that various proposed payment schemes can be analyzed. This could include SRBs, but also changes to the level of Basic Pay, changes to the retirement system, and changes to promotion rates. The basic idea is to calculate the annualized value of the cost of leaving the service in the present year and for each year in the future, determine the ACOL for the optimal time for leaving (when ACOL is maximized), and determine how that relates to the current reenlistment decision. This will be based on four general forms of present and future compensation:

- Regular military compensation (basic pay, housing and subsistence allowances, and the tax advantage)
- Bonuses (e.g., SRBs)
- Military retirement pay
- Civilian earnings

All future values must be predicted or derived in some way. While the Regular Military Compensation (RMC) components and military retirement pay can be estimated fairly accurately, it is quite difficult to estimate future bonuses and future civilian-earnings potential when service members leave the military.

The basic steps of the ACOL model are as follows:

(1) Calculate the present-discounted-annualized value of the stream of military compensation and civilian compensation for staying in the military for each possible leaving point, present and future. The difference in that annualized value between staying and leaving is the annualized-cost-of-leaving (ACOL) for a particular leaving point.

(2) Theorize that a service member will currently remain in the military (e.g., reenlist) if there is at least one ACOL value presently or at some leaving point in the future in

which the ACOL exceeds the service member's monetized value of his/her preference for military vs. civilian life.

(3) Insert the maximized-ACOL value into the reenlistment model, which is typically a logit model.

(4) To estimate the effects of a change in the payment scheme (e.g., an increase in an SRB, changes to retirement pay, a separation bonus), calculate how the new scheme would affect ACOL, and multiply that value by the estimated marginal effect of the ACOL variable from Step #3.

Besides the uncertainty in some of the pay components, there are many uncertain elements in this approach, which could create bias. We discuss those in Chapter IV. H.

C. ACOL-2 MODELS

Black, Moffitt, and Warner (1990) and Mackin (1996) modified the ACOL model by allowing the service preference parameter to have both a permanent component (as in the ACOL model), and a component that changes over time. This modification was introduced because the ACOL models were dynamically (time)-inconsistent for any decisions beyond the first reenlistment point. Inconsistency arises if the "optimal" plan of action is no longer optimal as time passes and new information arrives (e.g., shocks). Given that those who have a preference favoring military-vs.-civilian life are more likely to reenlist, most people (if not all) would be predicted to reenlist in subsequent decisions.

D. DYNAMIC RETENTION MODELS

The main difference in the DRM, from the ACOL and ACOL-2 models, is that there is not a time horizon that maximizes the ACOL, but rather there is a probability-weighted distribution of the cost of leaving for all future leaving points. These probability weights are determined within the model, which is one of the reasons why the computation is so burdensome and why models are limited in how much information they can incorporate. It is for this reason that Hogan and Black (1991) claimed that DRMs are less tractable than other retention models for policy analysis. While the advances in computer power have expanded the DRM's capabilities, limitations remain, as are discussed in Chapter IV. In practice, the DRMs have not been estimated with all the compensation components that are typically used in ACOL models.

The model involves estimating a steady-state retention rate with a given compensation scheme. This is based on relating an estimate of the taste parameter with changes in compensation, with the compensation being (theoretically) based on the same military components and civilian pay used in the ACOL-variable construction. With a new compensation scheme, the compensation changes are added into the compensation variable, and a new steady-state retention rate is estimated. The model estimates the retention effect of the compensation change. In addition, the DRM can estimate "transitional dynamics," which indicates how long it takes to transition to the new steady state.

The DRM was initially applied to military retention models to analyze the effects of "a broad range of compensation, retirement, and personnel policies" on retention (Gotz & McCall 1984). The model accumulates all military and civilian compensation into one variable so that the effects of any compensation change (from the present to a lifetime retirement annuity) can, theoretically, be estimated. However, as with the ACOL model, this depends on a host of assumptions, which we discuss in Chapter V. D.

The appendix contains a technical description of the DRM model.

IV. EVALUATION OF MODELS

A. GENERAL BIASES ASSOCIATED WITH MODELS ON RETENTION BONUSES

Models estimating the impact of bonuses on reenlistment and officer retention behavior have been used for two different purposes: estimating the longer term impacts of policy changes on retention behavior and estimating the short-term bonus required at a particular time to meet specific reenlistment or retention objectives. There are several general biases to these models that estimate the effects of SRBs. The same biases would apply to any model estimating the effects of bonuses. These biases speak to how difficult it is to apply estimates produced from such models to determine the optimal mid-career bonus for a particular group of service members at a particular point in time.

Successful estimation of a treatment effect in quantitative research requires that no other factors change as the treatment changes. To estimate the true causal effects of bonuses on retention rates, we would need one of the following:

- all other factors determining retention rates should be the same as the bonus varies over time, or
- these other factors that change are adequately controlled for.

This means that as the bonus changes in the model, no other factors relevant to reenlistment change. The first condition would be difficult to achieve. Various factors cause the changes in the retention rates that require changes in bonuses. So, the issue becomes whether those factors can be adequately controlled for.

The ideal situation for evaluating any policy or treatment would be through an experimental design, in which the treatment/policy were randomly assigned to people. Furthermore, the outcome would ideally be the intended behavior of the subjects rather than a decision made under constraints—that is, a limited number of slots for reenlistments. Unfortunately, neither of these is the case with manpower models, which makes it very difficult, if not impossible, to accurately estimate the causal effects of various forms of compensation.

This Chapter describes several biases that occur when estimating the effects of SRBs, as described in Arkes (2016). The biases described would also apply to estimating the effects of any compensation scheme.

The four major biases are as follows:

- Direct reverse causality—the bonus is being determined by the reenlistment rate.
- Indirect reverse causality—how the bonus is coded (i.e., the timing) depends on the reenlistment decision of the sailor.
- Measurement error—the bonus assigned to a reenlistment decision may not be the one that was considered at the time of the decision.
- Omitted variables bias due to excess supply—researchers may not observe all who are willing to reenlist with a certain policy (or bonus), as there may be constrained demand, leading to excess supply.

B. DIRECT REVERSE CAUSALITY

A regression model indicates how two variables move together, holding other factors constant. The regression model does not indicate why the variables move (or do not move) together, whether it is because the key explanatory variable affects the outcome, the outcome affects the key explanatory variable, they have a common factor that creates a correlation, or they just happen to move together by chance. An example of the latter is the annual numbers of Nicolas Cage movies and U.S. drowning deaths, which have a remarkably high positive correlation between 1999 and 2009. They are likely just incidentally correlated, as Nicolas Cage movies are not so bad that they'd cause drownings (or vice versa).[1]

Ideally, it would be random (or experimental) how SRBs are determined and assigned, and researchers would observe how reenlistment rates react to the bonuses. This would mean that the variation in SRBs across service members would be "good variation" in that it was random to the service members and would not be correlated with other things that could affect service members retention. But, we do not live in an experimental world, so it is likely that SRBs are set higher when retention is lower—this is "bad variation." This would contribute negatively to any relationship between SRBs and reenlistment rates. Assuming the true SRB effect on retention is positive, the reverse causality would understate the causal effect.

[1] See www.tylervigen.com/view_correlation?id=359.

This is demonstrated in Figure 1. We are interested in **A**, which is how the SRB affects the reenlistment rate. This effect is presumably positive. But the SRB would tend to decrease as the reenlistment rate increases, which is effect **B** (which is likely negative). This is one source of bad variation in the SRB variable. The estimated effect of SRB in any retention model would capture **B**, thereby leading to an understatement of the causal effect of the SRB on the reenlistment rate of **A**. The only way to correct for this is to, somehow, isolate the good variation in the SRB variable.

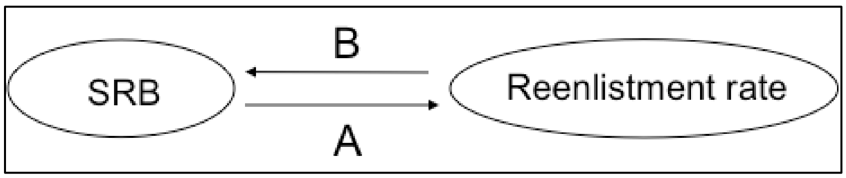

Figure 1. The Logic behind Reverse Causality

The standard approach to addressing this problem has been to include factors that could affect both the bonus and reenlistment/retention rates, including occupation (skill, in the Navy), year fixed effects, and perhaps variables representing the strength of the economy, such as the unemployment rate. But, the separate skill and year fixed effects still rely on the assumption that reenlistment propensity (regardless of SRBs) ebbs and flows similarly for all skills. Furthermore, factors specific to the civilian economic opportunities for those in a skill are not adequately controlled for, nor are the current military-environmental factors that could contribute to changing reenlistment rates. Thus, models typically do not adequately control for the things that affect retention and change with the bonuses. Other widely used methods to mitigate the reverse causality problem include the instrumental variables (IV) approach and quasi-experimental methods such as regression discontinuity (RD).

Some argue that this is not exactly reverse causality, as one service member's reenlistment decision would not affect the actual SRB. That is technically correct, but it is something very closely tied to the reenlistment decision (the general propensity to reenlist) that affects the SRB. Thus, we can also think of this as omitted-variables bias, with the omitted factor being "general propensity to reenlist," with a higher general propensity to reenlist leading to higher reenlistment rates and lower bonuses (thus contributing negatively to the SRB-reenlistment relationship).

C. INDIRECT REVERSE CAUSALITY

A second form of reverse causality occurs based on how researchers typically code SRB levels. Except for Arkes (2016), we believe that all studies used the SRB available for a sailor in the given skill at the sailor's decision date. The decision date is either the loss date or the reenlistment date. The loss date is typically the end of the active obligated service (EAOS), usually at the end of 36, 48, 60, or 72 months. The problem is that the reenlistment date is almost always different from the EAOS. The reenlistment date usually precedes the EAOS because people tend to reenlist early. But some sailors extend their service to postpone their decision to stay or leave.

A sailor intent on reenlisting can often choose when to reenlist. If the SRB changes over the course of the sailor's reenlistment window, he/she can attempt to reenlist when the SRB is higher. A Navy rule that ended in FY2008 allowed sailors 30 days after an SRB decrease to take the higher, pre-change SRB. While this is no longer the rule, sailors can guess when the SRB is relatively high for their skill. Sailors can extend to hold off their reenlistment decision if they believe the SRB is low. In contrast, sailors intent on leaving will not time their decision to a period of higher SRBs.

As an example, more nuclear-qualified service members (Nukes) reenlisted than the Navy had expected during the financial crisis of FY2009. This caused the SRB funds to run out by May of that year. The story from Navy manpower personnel is that no Nukes reenlisted for the rest of the fiscal year. Instead, they waited for the next reenlistment window or they extended so as not to miss out on the roughly $80,000 bonus. Arkes (2016) presented a notional example of this situation, shown in Figure 2. This is a case that could occur in a world in which the SRB has no effect on the probability of reenlistment. However, there would be a very high positive correlation between the SRB and the reenlistment rate.

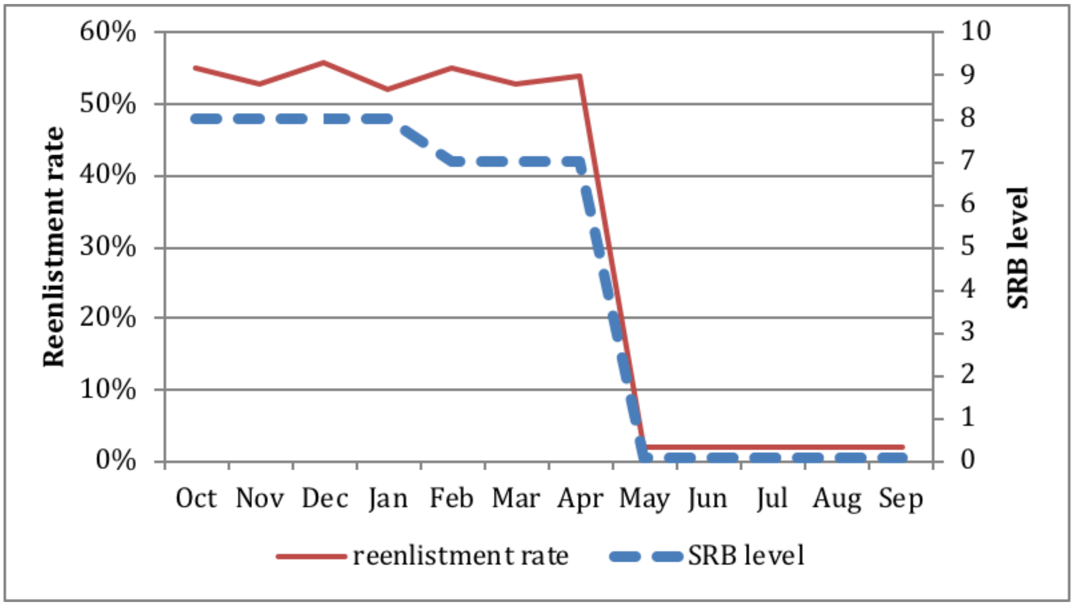

Figure 2. Notional Example of SRBs and Reenlistment Rates for Navy Nukes, FY2009.
Source: Arkes (2016).

A retention model is problematic only if the researchers code the SRB according to the eventual decision date instead of a fixed reference point, such as the ETS date. This would be indirect reverse causality because it is not that the sailor's decision affects the SRB level, but rather the sailor's decision affects how the researcher coded the SRB. The sailors who reenlisted were assigned a date that was more likely to be in a higher-SRB period. This indirect reverse causality would contribute towards a positive bias on the estimated SRB effect.

One of the shortcomings of many prior studies is that they were sparse on some of the key details. In particular, most studies neglected to indicate whether the SRB was coded according to the original ETS or the decision date. Given the lack of detail, it is a reasonable assumption that they just used the eventual decision date.

D. MEASUREMENT ERROR

Measurement error is typically described as a "fat finger" leading to miscoding. But measurement error can also occur when the variable poorly represents the concept or actual value being considered by the person. This can be the case with SRBs. A sailor up for reenlistment may consider multiple SRBs in their reenlistment. Yet, only one can be coded in a regression model. The sailor may decide based on the first SRB, but he/she may get coded a later SRB.

Furthermore, there could be cases in which a service member would be willing to reenlist without an SRB, but the sailor extends to reenlist when there is a positive SRB. As just described, a service member in such a case would typically be assigned the SRB at the time the eventual reenlistment occurs.

There are likely also cases in which a service member signs the contract to reenlist for a given SRB level, and the SRB changes in the few weeks it takes for the reenlistment to officially occur. The prior SRB would be the official one the service member received and considered for reenlistment, but the one at the reenlistment date would be the one the research records.

Measurement error typically results in a bias towards zero. Given that estimated SRB effects are typically positive, the likely bias would be negative.

E. EXAMPLE OF A BIAS FROM EXCESS SUPPLY

Consider Figure 3, which demonstrates a notional SRB-reenlistment relationship in a supply-demand framework. The quantity in this scenario is the reenlistment rate, and the price of reenlistment is the SRB multiplier. The demand curve (D) is the relationship between the Navy's desired reenlistment rate and the price the Navy would pay for that rate. For simplicity, we assume that the Navy needs a 50% reenlistment rate and would pay whatever SRB is needed to achieve that. The actual supply curve (S) indicates the relationship between the percentage of sailors who were willing to reenlist and changes in the SRB level. Note that the response variable (the reenlistment rate) is on the X-axis instead of its usual place on the Y-axis, which is done to be consistent with having the price on the vertical axis in a demand-supply model. We also include the dashed supply curve, indicating what the Navy believes to be the supply curve. In this case, the Navy believes it must pay more than is needed to achieve a 50% reenlistment rate.

The actual effect of an increase in the SRB by one multiple on the willingness of service members to reenlist is five percentage points, or 0.05. This can be calculated by taking two points on the actual supply curve and calculating [change in reenlistment-rate/change in SRB]. But, this 0.05 effect may not be identified with observed data because of a potential excess-supply problem.

Consider a case in which there are 600 sailors up for reenlistment in a given occupation and the Navy wants to retain 300 (50%) of them (see Figure 3). The Navy initially offers an SRB of 4, based on its belief of the supply curve. But, more sailors would reenlist than the Navy had planned at that SRB. The Navy would adjust at some point. Under the plausible scenario that the Navy keeps the SRB at 4 until they achieve their reenlistment goal of 300, they would then reduce the SRB level to 0 and they would not accept any more reenlistments. Thus, two points would be observed:

- 60% reenlistment rate at SRB = 4
- 0% reenlistment rate at SRB = 0

Calculating [change in reenlistment rate/change in SRB] again would produce an SRB effect of 15 percentage points—an exaggeration of the true 5-percentage-point effect. The source of the exaggeration is that the willingness of sailors to reenlist is not being observed, but rather whether they do reenlist. At the SRB of 0, more sailors were willing to reenlist than the Navy was accepting. Thus, there was excess supply.

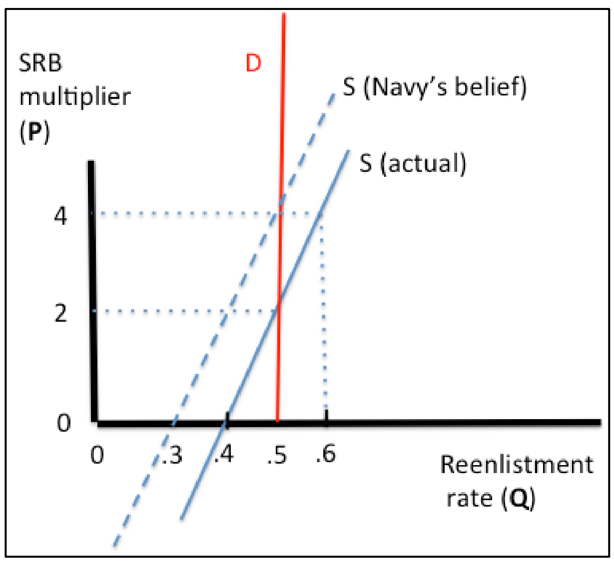

Figure 3. Notional Example of the Market for Reenlistments and Excess Supply

Consider perhaps a more plausible scenario. A Navy rating is being eliminated so no one can reenlist into that rating and the SRB is set at zero. The observed reenlistment rate would be zero for that rating even though some would have reenlisted if they could. If these non-reenlistments were compared to reenlistments from a positive-SRB period, then again, the SRB effect would be exaggerated.

Other scenarios could understate the SRB effect. Suppose that the Navy only allows 50% of sailors up for reenlistment to actually reenlist each month. If the Navy sets the SRB at 4, 60% of sailors up for reenlistment that month would be willing to reenlist, but the Navy would only accept 50% of them. The Navy may eventually realize that it set the SRB too high and reduce it from 4 to 2, and again obtain a 50% reenlistment rate. The SRB effect would be based on two points:

- 50% reenlistment rate at SRB = 4
- 50% reenlistment rate at SRB = 2

giving an estimated SRB effect of zero.

The bottom line from these examples is that an excess supply of sailors willing to reenlist—i.e., a constrained Navy demand—makes it impossible to accurately estimate the effects of any bonus, pay, or compensation policy change on reenlistment rates. The main cause is that researchers observe whether someone reenlists, not whether they were willing to reenlist. In the case of bonuses, the estimated effect of the bonus would be exaggerated if any excess supply occurs at a low SRB, while the estimated effect would be understated if the excess supply occurred at a high bonus level.

F. MINOR BIASES

Another source of bias is from MOS- (or rating)-switchers. Those who switch to a new MOS are almost certainly going to reenlist. This is a problem because they will tend to switch to higher-bonus MOSs. This creates an upward bias to the bonus effects. However, keeping switchers with their original MOS would result in a downward bias, as it would seem like more service members are willing to reenlist at a low-or-zero bonus.

One other source of bias comes from attrition before any reenlistment decision. If there is greater attrition among a cohort for some reason, it may mean that many of the "poor fits" have left and more of the remaining service members would reenlist. For attriters, the SRB would probably have a minimal effect compared to those not attriting. Thus, the greater the number of attriters, the more exaggerated the SRB effect would be.

G. SUMMARY OF THE MAIN BIASES

Table 1 summarizes the four main biases with a short description and the likely direction. The opposing biases, uncertain sizes of the bias, and even uncertain direction of the bias from "excess supply" means that there is no way to determine if the cumulative effects of these four biases is positive or negative.

Table 1. Summary of the Main Biases in Bonus-Reenlistment Studies

Bias	Description	Likely direction
Direct reverse causality	The propensity of sailors to reenlist likely affects the SRB level.	Negative
Indirect reverse causality	Researchers code the SRB at the loss date or the time of reenlistment rather than using a fixed time.	Positive
Measurement error	The SRB coded may not reflect the SRB considered at the time of the decision.	Negative
Excess supply	Researchers observe whether the sailor reenlisted, not whether the sailor was willing to reenlist.	Positive or negative

H. CAN THESE BIASES BE ADDRESSED?

Theoretically, the first two biases could be addressed, but any correction potentially exacerbates other problems. Arkes (2016) corrected for the indirect reverse causality by using the SRB based on the original ETS—this means that the value of SRB assigned to an individual does not depend on whether and the timing of their decision. But, this could potentially increase measurement error.

Arkes (2016) attempted to address the direct reverse causality problem by using skill-FY-interacted fixed effects, so service members were only compared to other service members who were in the same skill and who had an ETS in the same fiscal year. This means that service members were just compared to others who had similar factors that would affect general reenlistment propensity (or supply), such as the civilian-sector demand for the skills of sailors in a given Navy skill and the Navy climate (e.g., OPTEMPO) for those in the given skill. Thus, the

SRB effect on reenlistment is identified by within-skill-FY variation in SRBs and reenlistment decisions. This is more likely to be *good variation* because it largely comes from Navy mid-year corrections to SRBs based on mis-calculations of too many or too few service members reenlisting at the given SRB.

There are two downsides to using skill-FY-interacted fixed effects. First, it changes the relative weights in the model for each skill and each FY. When fixed effects are applied, the overall estimate of SRB is the average of the within-group SRB effects, weighted naturally by the percentage of observations. That is, the overall estimated SRB effect is[2]

$$\beta \quad = \sum_{g} \beta_g * \left[\frac{\Pr(g) * var(SRB|g)}{\sum_g [\Pr(g) * var(SRB|g)]}\right]$$

where β_g is the SRB effect for group g, Pr(g) is the percentage of the sample in group g, and var(SRB|g) is the within-group variance of SRB. The latter is the key, as some groups will have significant across-FY variation in SRBs but never any within-FY changes in SRBs. And so, the relative weights each skill-FY group has in its contribution to the overall effect changes when going from separate skill and FY fixed effects to skill-FY-interacted fixed effects.

The second problem with this fix for direct reverse causality is that it could create more of a (likely downward) bias from measurement error. The extent of any bias from measurement error is based on the percentage of the relevant variation in the explanatory variable that is subject to measurement error. The fixed effects will reduce a lot of the normal variation, leaving the variation from measurement error as a larger percentage.

The bottom line is that eliminating these biases is almost certainly impossible, so any study on estimating the effects of bonuses should note that such biases remain.

There is no way to know the extent of any bias from excess supply. While some argue that measurement error can be addressed, such methods are rife with questionable assumptions.

---

[2] This interpretation was formulated in Gibbons and Suarez Serrato (2011), Angrist and Krueger (1999), Wooldridge (2005), and Angrist and Pischke (2009).

V. ASSESSMENT OF ACOL MODELS AND THE DRM

In this chapter, we highlight several data and modeling issues that that were identified when reviewing the DRM and ACOL models. These empirical problems are in addition to the ones mentioned in the previous chapter. Thus, they expand the list of data and modeling issues that will hinder future efforts to predict retention decisions. The empirical problems include the following:

A. Omitted-variables bias where the compensation variable effect picks up the effect of all the other retention factors that come with it (e.g., September 11th, the Financial Crisis, different MOS conditions across service members)

B. Too much *bad variation* in the compensation variable coming from the highly-problematic bonus component

C. A catch-22 for what to do with year and MOS effects

D. Significant reliance on indeterminate assumptions underlying the models

E. Unstable estimates with regards to seemingly minor modeling decisions

F. Issues with extrapolating to the mid-career bonus and estimating retirement effects

We use the term *occupation* or *MOS* (rather than *skill*), as most dynamic-programming models have examined all services.

A. OMITTED-VARIABLES BIAS

As mentioned in the prior sub-section, any credible retention model estimation would use variation in compensation that is independent of variation in other factors that could affect retention. That is, if we were to compare two service members, the only difference between them in terms of retention factors would be the compensation they receive. This means that the military environment (e.g., enjoyment of their military position, OPTEMPO), civilian opportunities, and other retention factors would need to be equivalent. We would then see how the higher compensation for one of them is associated with higher retention rates.

The main problem is that there is usually a reason why compensation is different across people. That reason could have its own effects on retention. Using the idea on bonuses from the prior chapter, higher bonuses are usually associated with a lower propensity for people to reenlist (perhaps due to better civilian opportunities or an unpopular war).

Most variation in compensation would come across years of service, rank, fiscal year, and MOS (with regard to bonuses). To the extent that these factors are then not controlled for, the compensation variable/component of the model would capture the effects of those factors.

The existing DRM studies have not controlled for any year or MOS effects (except when the analysis focused on a particular MOS), perhaps due to limitations of computing power. The ACOL studies sometimes controlled for year and sometimes controlled for MOS, but not the interacted year-MOS, as in Arkes (2016). Also, they did not incorporate SRBs and bonuses into the military compensation calculations. Without controls for the year, the variation in the pay components largely comes from across-year variation. This means that the estimated effects of total compensation (or the corresponding variable to the ACOL variable) would capture the net effects of everything that affects reenlistment rates in a given year. These would include the following:

- civilian economic opportunities not captured in the civilian-income variable used
- the military environment, such as OPTEMPO, the intensity of deployments, the command climate, the attitude towards the military, etc.
- any incentives, including bonuses and non-monetary incentives

The period effects are being attributed to the compensation differences. This is biasing the compensation effects.

Imagine the scenario in which the Navy wants to maintain a 50% reenlistment rate every year. In some years, say due to high OPTEMPO, more than 50% of the sailors with ETS at the beginning of the year may choose to leave military service. In response, the Navy could increase the bonus and eventually they obtain the 50% reenlistment rate. In another year, suppose that there is a deep recession, and so more sailors want to reenlist. The Navy does not have to offer any bonuses that year, as they can easily achieve the 50% reenlistment goal—demand constraints would prevent them from allowing all willing sailors to actually reenlist. Thus, every year they get the same 50%

reenlistment rate, but the general compensation level would vary. The estimated effect of compensation would then be zero because it would capture the effects of each factor (OPTEMPO, civilian opportunities, bonuses, and demand constraints) on the reenlistment rate, the effects of which would sum to zero each year, in this scenario.

The problems of making across-time comparisons was demonstrated by Arkes (2017). In a model meant to demonstrate such problems, he estimated a time-series model of U.S. gross domestic prodcut growth rates on the highest marginal income tax rate from 1990–2015. He found a significant positive coefficient estimate on the tax rate—giving the counterintuitive result that higher taxes were associated with stronger economic growth. The simple explanation for this result is that the variation in tax rates was picking up the effects of other things that changed concurrently. In particular, President Clinton and Congress raised tax rates right before the internet boom led to the strong growth of the mid- to late-1990s. President Bush (and Congress) then lowered the tax rate in 2001, and those lower taxes coincided with the recessions associated with the dot-com bust, the September 11 attacks, and the financial crisis. Without controlling for the time period (which a time-series model can't control for), the estimated effect of the tax rate picked up the effects of other factors affecting economic growth.

To apply the same reasoning, not controlling for MOS (or rating/skill), means that the effects of differences in compensation across MOSs (from RMC, SRBs, or civilian pay) would capture the effects of all other factors affecting those MOSs. For example, an MOS that has great civilian employment opportunities would have low reenlistment rates because of those opportunities. The higher expected SRBs for that MOS would reflect the effects of the civilian opportunities for that MOS.

Hansen and Wenger (2002) demonstrate these problems (and subsequent problems with the ACOL model we discuss in Chapter VI. B.) by showing how the estimated pay elasticity and SRB effect on reenlistment rates depend on various assumptions and modeling decisions. Table 2 presents the key results from their analysis. The baseline model had an estimated pay elasticity of 1.5% (a 1% increase in pay increases reenlistment rates by 1.5%) and a 2.5-percentage-point SRB effect on reenlistment rates. When they add in rating controls, the estimates increase substantially.

This suggests the presence of negative bias from omitted-variables bias (or reverse causality) without controlling for ratings. They also find that going from the baseline of including fiscal-year controls to excluding them results in lower estimates, again perhaps a negative bias from omitted-variables bias from some incidental set of circumstances.

Table 2. ACOL-Model Predicted Pay Elasticities and SRB Effects with Various Modeling Decisions.

Adapted from Hansen & Wenger, 2002.

	Estimated pay elasticity	Estimated effect of one level of SRB multiplier
Baseline model (assumes 20% discount rate, contains fiscal-year controls, but not rating controls)	1.5%	2.5 pct. pts.
Change from baseline:		
Exclude fiscal-year controls	1.0	1.6
Add rating controls	2.7	4.4
10% discount rate	1.5	0.9
30% discount rate	1.5	3.3
Separated military and civilian pay	2.8	1.9
Exclude race/ethnicity and age	0.9	1.5

B. TOO MUCH "BAD VARIATION" IN THE COMPENSATION VARIABLE COMING FROM THE HIGHLY-PROBLEMATIC BONUS COMPONENT

From the components of pay that are used in the calculation of the compensation variable, the bonus (e.g., SRB for enlisted personnel) is, by far, the source of the most variation. The other components (RMC, retirement pay, and derived/estimated civilian earnings) are fairly stable across people and over time. But, bonuses can be significant for some (e.g., currently around $90,000 for enlisted Nukes and $100,000 for Surface Warfare Officers) and zero for others.

Having bonuses as the primary source of variation for the compensation variable is problematic because of the reason mentioned above; bonuses are typically set higher when service members' propensity to reenlist is lower (in addition to the other sources of bias, including measurement error). Thus, the estimated effects of the compensation component should have just as much bias as the estimated effects of the SRB variable in the simple model described in Chapter IV. H. These biased estimates would then be applied to changes in ACOL from retirement policy changes or other changes in compensation schemes.

From what we could elicit from the reports, ACOL models typically include bonuses (SRB or officer bonus) in the compensation variable, but the DRM typically has excluded the bonus. Not including the bonus could result in omitted-variables bias (as mentioned above), while including bonuses would make them the largest component of variation in the compensation variable; however, bonuses systematically pick up "other factors of retention" (e.g., all the reasons why the Navy would have to increase the bonus in the first place). This creates even greater omitted-variables bias. Either way, there is omitted-variables bias.

C. THE CATCH-22

There's a catch-22 here. These points indicate that any analysis of the effects of SRBs or pay on the probability of service members reenlisting should hold constant the year and the MOS (or skill) to avoid omitted-variables bias from the SRB/pay picking up the effects of other things associated with the MOS or the year. However, controlling for differences across years and MOSs means the remaining variation is almost entirely from SRBs, which are highly subject to reverse causality and omitted-variables bias.

The bottom line is that estimating the future impacts of changes in bonuses or pay is a complex modeling challenge. It is not just uncertainty in the estimates due to wide standard errors; there are systematic biases that almost certainly move the estimates significantly away from the true SRB/pay effects. This highlights the challenge in estimating the effects of the new retirement system or other changes in compensation.

D. SIGNIFICANT RELIANCE ON INDETERMINATE ASSUMPTIONS UNDERLYING THE MODELS

The DRM and ACOL models rely on a series of critical assumptions. Some of these assumptions could be relaxed, but several are integral to the models. Here are just a few of the assumptions that significantly influence the model estimates:

- The model assumes a discount rate so that future payments can be put in present-value terms. The models incorporate an estimated discount rate and assume that it is the same for all service members and constant over time. The discount rate has, at times, been estimated within the model, but this requires another set of assumptions.
- The calculations for many components of compensation are also estimates, particularly future bonuses and civilian pay when service members separate from the military. This presents two major problems. First, basic economic theory suggests that, due to general risk aversion, more value is placed on $1 of certain income than $1 of uncertain income. But, certain and uncertain income are treated the same in these models. This uncertainty introduces measurement error in the ACOL or compensation variable. This would add to the already existing bias from measurement error in the bonus variable described above.
- The models assume that service members know how long they will be alive and collecting retirement pay.
- The models assume that service members know the distribution of the *military service preferences* shocks as well as their current values.
- The models assume that taste parameters for the military and civilian life are constant throughout life.
- The models assume that service members have linear utility functions; this allows their income to be discounted rather than discounting the nebulous concept of their utility.

Most of these assumptions are naturally part of the model, and so we cannot calculate exactly how much the model's estimates depend on the assumptions. But, some can be tested for. Hansen and Wenger (2002) demonstrate, as shown back in Table 2, that a lower discount rate (from the baseline rate of 30% to the changed rate of 10%) lowers the SRB effect from 3.3 to 0.9 percentage points. The 30% discount rate is probably unrealistically high, but nevertheless results in a higher SRB effect. This reflects that the present gets weighted much more than the future, so current SRB payments would have a greater effect on retention. This demonstrates how important basic assumptions are to the models' results.

Further, Table 2 shows that the pay elasticity almost doubles and the SRB effect is cut by about one-third when military and civilian pay is separated. For the pay elasticity, this may reflect measurement errors associated with civilian pay causing a bias towards zero. This shows it is important to separate military and civilian pay (as the simple model automatically does by having a variable for SRB).

E. UNSTABLE ESTIMATES WITH REGARDS TO SEEMINGLY MINOR MODELING DECISIONS

Simple modeling changes that should have minimal effects on the estimates end up having large effects, suggesting an instability in the model estimates. From Table 2, Hansen and Wenger (2002) find that simply excluding race/ethnicity and age results in a 40% reduction in both the pay elasticity and SRB effects. This change should theoretically have minimal effects, as race/ethnicity should not be correlated with the SRB or military pay. This may just come from variation in civilian pay.

F. ISSUES WITH EXTRAPOLATING TO THE MID-CAREER BONUS AND ESTIMATING RETIREMENT EFFECTS

There is another issue specific to estimating the effects of retirement changes. The perceived retirement annuity period stretches from around 10 to 50 years in the future (from military retirement to expected end-of-life). The present value of these payments depends significantly on the discount rate. Any slight error in the discount rates service members are assumed to use could cause major errors in calculating the future value of both military and civilian compensation, including retirement benefits, and so we can't credibly calculate the retention effects from changes to the retirement annuity.

Furthermore, any retention model would typically be based on the prior five to 15 years of data. Using those historical patterns to make inferences on what will occur in the next two to 20 years is highly suspect. Any structural changes to the force would bias model estimates. For example, the shift towards more technical occupations could cause future service members to respond differently to compensation changes from how service members have historically responded.

G. SUMMARY

The biases described here highlight the data and modeling challenges in estimating compensation effects in any retention model. Not only are there unavoidable errors, there are systematic biases. In Chapter VI, we describe the conditions needed to obtain credible estimates of the retention effects of retirement-scheme changes and mid-career bonuses. We also discuss how we can improve the existing models.

VI. DESCRIPTION OF KEY STUDIES

A. SELECTED STUDIES USING THE BASIC RETENTION MODEL

Hattiangadi, Ackerman, Kimble, & Quester (2004)

In this Center for Naval Analysis (CNA) report, Hattiangadi et al. (2004) focused on estimating the effects of the new Marine Corps policy, starting in 2000, that paid SRBs as up-front lump sums, rather than the original method of one-half up front and the other half in annual installments over the course of the reenlistment. Thus, they regressed whether a Marine reenlisted on the SRB level, an indicator for a Marine being under the lump-sum policy (a positive SRB and being up for reenlistment in the 2000–2003 period), and other controls.

This was a basic model (as described in Chapter III. A.). The authors attempted to address the potential direct reverse causality problem with MOS controls and monthly (national) unemployment rates and military/civilian wage rates. Still, underlying changes in reenlistment propensities over time within an MOS likely led to changes in the SRB, which leads to the direct reverse causality.

Furthermore, the authors used the SRB at the decision point, which, for reenlisters, could be timed strategically. This should lead to an upward bias in the estimated SRB effect from indirect reverse causality.

The estimated "lump-sum-policy" effect may also be biased. Without year controls, the estimate may reflect the effects of any factors specific to the FY2000–2003 period relative to the FY1985–1991 and FY1998–1999 period—the drawdown years of FY1992–1997 were separately controlled for. These FY2000–2003 factors could include the September 11 attacks, the dot-com boom and subsequent bust, the start of the Iraq War, and more. It is likely a positive bias, as the estimated effects of the lump-sum policy were much higher than would be expected, implying an unfathomable discount rate for Zone A potential-reenlisters of 155%.

The estimated effects from a one-level increase in SRBs were 6.6 percentage points for Zone A reenlistment decisions, 7.2 percentage points for Zone B decisions, and

3.5 percentage points for Zone C decisions. The first two seem much higher than would be realistic, perhaps a product of a positive bias from indirect reverse causality.

Hosek & Martorell (2009)

This study (Hosek & Martorell, 2009) was intended to estimate the effects of non-hostile and hostile deployments on the probability of reenlistment for all four services. The authors argued that SRBs represent a potential confounding factor, so they need to be controlled for (and the effects of which need to be accurately estimated). They gave extra attention to the problem of direct reverse causality. However, the only approach to address this was to include occupation fixed effects.

Hosek and Martorell's (2009) calculation of the SRB variable made their study more susceptible to the positive indirect reverse causality bias. They did not have actual SRB levels, so they calculated the average SRB level for service members from a given occupation with a reenlistment decision in a given quarter. The concern, again, is that those intent on reenlisting will do so when the SRB is higher, or perhaps delay their decision if the SRB is too low. As with other studies, there would be positive and negative biases on their estimates, and it is uncertain which dominate.

Asch et al. (2010)

As part of a larger study on how cash incentives affect various military manpower outcomes, Asch et al. (2010) examined the effects of SRBs for first- and second-term reenlistment decisions for all four services. They used the same method of calculating SRBs as Hosek and Martorell (2009), again creating a positive bias from indirect reverse causality. The other biases mentioned in Chapter IV also apply. Asch et al. (2010) estimated SRB effects over the FY2002–2007 period for first- and second-term decisions for each service. And, they estimated models with and without controlling for whether the service member was on deployment at the time of the reenlistment decision, which only meaningfully mattered for Army reenlistment decisions. The estimated SRB effects ranged from being negative and significant (Navy, second-term decisions) to as high as 0.089 (for Army, first-term decisions, when SRB is allowed to vary based on deployment).

Arkes (2016)

In this study, the author examines the effects of SRB level on Navy first-term reenlistments for those with four-year initial obligations. As described in Chapter IV. H., Arkes (2016) attempts to correct for direct and indirect reverse causality but acknowledges that he is not able to account for measurement error and excess supply biases. Nevertheless, he estimates that a one-level increase in the SRB leads to a 3.6-percentage-point increase in the reenlistment rate.

B. SELECTED ACOL STUDIES

Hansen & Wenger (2002)

In this CNA report, the authors estimated pay elasticities and SRB effects under different assumptions used in the ACOL model. The authors did not indicate all the assumptions and coding strategies used (e.g., how they coded SRBs), but it appears that the baseline model included FY controls but not occupation controls.

They found that the estimated SRB effect changed significantly when inputs or assumptions changed. For example, going from a 10% to a 30% discount rate increased the estimated SRB effect from 0.9 to 3.3 percentage points. This highlights the sensitivity to discount rate assumptions and the advantage of more basic models that do not rely on discount rate assumptions or taste parameters.

A published version of this report (Hansen & Wenger, 2005) just focused on pay elasticities and did not cover SRB effects, thus it is not as relevant for our purposes.

Warner & Goldberg (1984)

Warner and Goldberg (1984) studied the effect of sea duty (i.e., a non-pecuniary factor) on the reenlistment decisions of enlisted service members in the U.S. Navy. They first derived a theoretical model of reenlistment decisions that predicted that sea duty negatively affects retention. Then they evaluated that hypothesis using actual data on reenlistment decisions and found that, indeed, retention was inversely related to the incidence of sea duty.

C. SELECTED DRM STUDIES

DRM, which was developed at RAND, has been extended in various ways to tackle a myriad of other topics in military manpower policy. Asch, Johnson, and Warner (1998) and Asch and Warner (2001) analyzed how changes to the retirement benefit system and basic pay would impact retention. The former forecasted the impacts of changing the military retirement system to one resembling the Federal Employees Retirement System (FERS). The authors found that such a conversion, without progressively higher pay raises (by rank) and more generous retention and separation bonuses, would lead to undesirable seniority profiles. In the latter paper, the authors greatly extended the theoretical model by adding individual ability and effort to the model. Ability, effort exertion, and noise impact promotion probability, which is a rank-ordering process. The model also included the potential for involuntary separation. The model was then solved for optimum effort by labor market participants and ability sorting by rank and retention response to compensation policy was simulated. In both papers, previously estimated parameters from DRM were used to run simulations. No econometric estimation was performed.

Hosek, et al. (2004) extended the model to include the initial decision to enlist, looking specifically at information technology (IT) workers in the military. During the height of the dot-com boom (and subsequent bubble burst), concerns were raised that the military would be unable to compete with the civilian IT sector in attracting and retaining sufficient personnel with the necessary technical skills. The analysis showed that the military had been successful in retaining such personnel and would continue to be able to do so in the future. DRM uses actual wage and retention data of IT personnel to show how the value of training in IT yields higher retention rates.

Mattock and Arkes (2007) examined retention for Air Force officers and the effectiveness of the Aviator Continuation Pay Program, incorporating civilian wage data for pilots. Results and simulations showed that while modest changes in the program yielded minimal changes in retention, up to 15% of experienced pilots may be lost if the bonus pay is eliminated altogether.

Asch, Mattock, and Hosek (2013) extended DRM to calculate retention cohort size as new policies were introduced and followed them through time, estimating the transition path to the new stable equilibrium. Asch, Mattock, and Hosek (2017) examined the potential impact of changes to the BRS across each of the services. Transitions (changes in the mix of individuals under the new and old systems over time), steady-state retention (the behavior of cohorts who spend their entire career under the new system), and costs under BRS were also estimated. The authors found that the new retirement system could maintain the same force size and mix as the legacy retirement system.[3]

The DRM reports have been notably sparse in discussing many details, but these details raise important questions for understanding the validity of the models, including the following:

- Did they calculate bonuses and, if so, how did they calculate the bonuses—based on actual or original ETS date?
- What was the major source of variation in pay that drove variation in the ACOL variable or the compensation variables?

[3] This is not an exhaustive list of extensions and applications of the original Gotz-McCall (1984) model, but it does represent a good cross-section of the ways in which the model has been pushed forward.

THIS PAGE INTENTIONALLY LEFT BLANK

VII. FUTURE EFFORTS TO OBTAIN CREDIBLE ESTIMATES FROM THE DYNAMIC PROGRAMMING MODELS

The RAND DRM studies Discussed in Chapter VI. C. did an admirable job in pushing data-driven decision-making forward, given constraints in data and computation power. They estimated sophisticated, internally-consistent dynamic models that incorporate rational forward-looking agents. The cost of such a model was that they had to take short-cuts and make sweeping assumptions to be tractable. This meant that they were unable to address several sources of biases that, unfortunately, were not fully acknowledged. These biases were a necessary cost of doing business at the time.

We aim to leverage better data, more computational power, and cutting-edge econometric techniques. We may not be able to solve all the problems and biases, but we should be able to mitigate a subset, and having full knowledge of potential remaining biases should inform policy-making.

We described earlier the issues associated with all retention models and issues specific to the dynamic models (ACOL and DRM). These issues need to be addressed for any model to produce estimates predicting the short-term and long-run effects of a compensation-policy change.[4] Again, it isn't just that there is unavoidable error, but rather that there is likely be a systematic bias.

Perhaps the most important issue is that the variation in compensation would have to be effectively random. The current reliance on across-year variation or across-MOS variation is almost certainly capturing the effects of unobserved year- or MOS-specific factors. But, as Arkes (2016) pointed out, even controlling for both year effects and MOS effects separately would not solve this problem, as the reenlistment propensity (based on civilian opportunities and the military environment) can ebb and flow differently across MOSs. Thus, there should be controls so that service members are only compared to others in their MOS in a short period of time, say a fiscal year.

[4] While the ACOL only can estimate long-run effects, the DRM can estimate both the long-term effects and the short-term effects during the transition. However, the accuracy of these estimates has never been tested.

Furthermore, the dynamic model we aim to develop would attempt to incorporate non-monetary benefits, consider and document the natural error in models, and account for the value of high-certainty versus low-certainty in future income (military and civilian).

Another important feature of the model we are developing is more accurate measures of a service members' civilian earnings potential. These are just guessed at quite crudely in existing ACOL and dynamic models. Matching IRS or Social Security data to personnel data for exiting sailors would constitute a significant improvement over the current methods. However, even this would be subject to significant measurement error.

Other conditions to obtain credible estimates would require rule changes or better data collection. For example, to avoid significant measurement error, a credible model would require bonus rules to be changed so that each service member up for reenlistment is only eligible for one possible bonus. If the bonus changes during a reenlistment window (or an extension provides opportunities for higher SRBs), then the researcher doesn't know which bonus the service member used to make his/her reenlistment decision. We fully understand that there are costs of having a more rigid bonus system.

We would also need extra data to avoid some biases. This includes personnel data indicating when the reenlistment became official, as the bonus depends on what was available at the time the contract was signed, not when the reenlistment began.

Demand constraints or excess supply also cause biases. Biases are reduced if everyone who wants to is able to reenlist or if we know whether they want to reenlist (or, for officers, to continue). Alternatively, personnel data would need to indicate which service members applied for reenlistment.

Other reenlistment rules that cause estimation problems include the following:

- being able to extend and taking a higher bonus when it becomes available
- switching skills and taking the bonus for the new skill
- allowing tax-free bonuses for those on deployment

- some high-tech skills having more than one reenlistment window (e.g., Nukes).

We are not arguing to change the reenlistment rules. Rather, we are pointing out that these rules create further challenges for estimating the effects of compensation on reenlistment decisions.

Beyond all the conditions needed, the credibility of model estimates would still depend on a host of assumptions:

- The single discount rate applied is not the actual discount rate considered by the whole sample
- The present-discounted-value of any dollar of compensation in the future does not depend on how uncertain that compensation is. That is, service members equally value $1 of fairly certain future income (e.g., basic pay) and pretty uncertain future income (e.g., future SRBs or civilian earnings after leaving the military).
- There is no systematic directional bias in expected income in the military or the civilian economy.
- A service member’s utility is linear with respect to any compensation.
- The service members know the distribution of the shocks (epsilons) as well as their current values.
- The taste parameters for the military and civilian life are constant throughout life.
- Historical patterns in how service members respond to compensation changes will continue in the future.

In sum, there will inevitably be biases in any retention model. Our ideas here are meant to minimize the bias to maximize the value of a dynamic-programming retention model.

THIS PAGE INTENTIONALLY LEFT BLANK

VIII. DATA

Our ability to predict individual behavior relies heavily on data. The suite of econometric models we propose will be most useful when paired with more detailed data. Indeed, some of the biases we have detailed throughout this report can be eliminated, or at least mitigated, with better data.

In this chapter, we outline our data “wish list.” We do not expect that a complete dataset that captures *all* the requirements specified here can be provided. Instead, the following description should inform what additional capabilities our new tools will be able to provide, given the appropriate data.

A. INDIVIDUAL SOCIO-ECONOMIC AND PROFESSIONAL CHARACTERISTICS

At the risk of belaboring the obvious, individuals’ labor market decisions depend on their socio-economic characteristics. Race, gender, age, marital status, dependents, financial status, academic background, and ability and personality all impact one’s decision to join, stay, and leave the military. Having as detailed a dataset of socio-economic *individual-level* characteristics will lead to more accurate retention predictions. Beyond aggregate numbers, having individual characteristics allows us to analyze the labor market decisions of population sub-groups that may be of interest to the Navy in the future.

In addition, an individual’s professional characteristics are expected to be important in our analysis. Data such as an individual’s MOS; where, when, in what capacity, and under whose command he/she served; and FITREP and other professional evaluations can all be valuable in informing an individual’s decision. Ad-hoc (retention) bonuses and other non-standard pay or working conditions experienced by different individuals may also be important in predicting retention. If retaining quality personnel (in addition to filling billets) matters for the Navy, augmenting basic demographics with data on performance will be useful in generating and testing policies that induce more highly capable individuals to stay longer.

B. PANEL DATA

Previous models that estimated retention probabilities purposefully used a very parsimonious set of data on the professional careers of individuals. In particular, the DRM only captured aggregate numbers of stayers and leavers by cohort. The model then treated everyone in the data as identical (except for "taste for military life") in two important ways:

1. The socio-economic and professional characteristics described above played no part in how individuals made labor market decisions.
2. Individuals *across time* were assumed to be identical when they clearly were not. DRM solves and estimates how a rational, forward-looking individual makes optimal decisions from the start of his/her military career until the end. The use of cross-sectional data in past studies, however, has meant that cohorts had been "stitched" together through time. For example, data of naval officers collected at a single point in time will have observations of O-1 through O-10 making career-decisions during that year. Without the ability to follow an officer through his/her career, the econometric model necessarily assumes that the O-1s in the data will behave exactly like the O-2s in the data after they are promoted. The O-2s will in turn behave exactly like the O-3s upon promotion, etc.[5]

It would be most beneficial to have a panel (or longitudinal) data that follows individuals throughout their military careers. While some characteristics (such as race, personality, etc.) are time-independent, others (such as professional evaluations, marital status, dependents, etc.) can and will change throughout an individual's military career. These changes can profoundly impact labor market decisions. Knowing when these changes occur will aid in accurately predicting retention behavior.

C. OUTSIDE CIVILIAN DATA

While professional careers within the military (pay, rank, grade, and other professional characteristics etc.) can feasibly be captured, previous models have not had the capacity to observe individuals once they chose to separate. Most models have used a rough average income from the civilian sector as a proxy for what soldiers and officers

---

[5] Again, we wish to emphasize that this was a feature of the DRM model that allowed the estimation of this complicated dynamic model using a very compact dataset given the computational constraints of its time period.

would earn after leaving the military. Understanding the “outside” option is vital to more accurately modeling and predicting retention.

We propose that data of soldiers and officers, once they separate, be linked to Veterans Administration (VA) and other civilian financial data to the extent possible. In particular, health data from the VA and income data from the Internal Revenue Service (IRS) or Social Security Administration (SSA) will be hugely beneficial in understanding why individuals may choose to leave (or stay) when they do.[6] The usefulness of this type of data cannot be overemphasized. Reforms to military compensation or the retirement system and their impacts on recruiting and retention must be considered in the context of the attractiveness of the civilian market to military personnel specifically.

[6] We anticipate that establishing links to civilian data, particularly IRS or SSA, will be difficult without significant buy-in (and influence) of Navy leadership.

THIS PAGE INTENTIONALLY LEFT BLANK

IX. ALTERNATIVE APPROACHES TO SETTING OPTIMAL BONUSES

Given the challenges any model faces to produce credible estimates, we suggest alternative approaches to augment models in determining the optimal bonus.

A. APPROACH 1: THE ROLE OF AUCTION MECHANISMS IN SETTING BONUSES

One goal for this and subsequent years is to determine the most efficient mid-career bonuses to offset the reduced retirement annuity incentive to stay in the Navy. This would help increase retention and retain the more talented service members. A retention model, be it a simple model or a more complex DRM, will naturally and unavoidably have errors in estimating the required bonus at a particular time for a particular service member or sub-group of service members due to the following:

- natural variation and randomness
- potential biases that cannot be addressed
- changing conditions (the civilian labor market, the military environment, the command climate) from the analysis period to the present when the suggested bonuses are being set

As a result, bonuses will generally be set inefficiently. Some will be too high, costing the Navy more than necessary. Some will be too low, requiring adjustments to meet retention needs. Furthermore, models do not currently incorporate personalized non-monetary retention incentives or address the quality of service members retained.

An auction mechanism may turn out to be the optimal method for determining the most efficient set of retention/reenlistment bonuses or other special and incentive (S&I) pays. In fact, the Quadrennial Review of Military Compensation (QRMC) concluded that “the Services should explore [S&I] pays, such as reenlistment bonuses, which could potentially use an auction mechanism to incorporate member preference into payment rates” (Under Secretary of Defense for Personnel and Readiness, 2008, p. 18).

The idea behind an auction mechanism is that service members up for reenlistment in a particular skill group would bid the bonus they would need to stay in the Navy. The auction can be designed to elicit truthful reservation prices, or the lowest incentive service members would accept to reenlist. If well designed, a retention auction should automatically indicate the lowest bonus necessary to attain the desired retention rate. Furthermore, bonus costs could be reduced beyond that, as the Navy could use location preferences and other non-monetary factors in conjunction with monetary bonuses. Finally, an auction could incorporate talent-related incentives to provide greater incentives to higher quality service members (Coughlan & Gates, 2010).

Auctions have several advantages over the complicated models previously described:

- They do not rely on the numerous assumptions that are necessary to estimate the various models, such as on the discount rate.
- They do not rely on the major assumption that service members would respond to current bonuses as they have responded historically, on average.
- They avoid the inherent biases that plague the existing models, such as reverse causality and unobserved heterogeneity from excess supply.
- They indicate what the actual minimum bonus would be, in contrast to the above models, which just give a best guess for the optimal bonus and would almost certainly be too high or too low.
- They can reduce the bonus by combining it with non-monetary incentives.

The 2008 QRMC (Under Secretary of Defense for Personnel and Readiness, 2008, p. 10) identified four attributes for evaluating military compensation programs. Adapting these four principles to the specific application of retention incentives, we generate the following performance measures (Coughlan & Gates, 2010):

- Voluntary: retention incentives should be structured such that each service member willingly serves and perceives that compensation for the assignment is both satisfactory and fair.
- Flexible and Responsive: retention incentives should be flexible enough to quickly and effectively adjust resources to respond to emerging issues, shifting priorities, and changing market conditions.
- Best Value: retention incentives should provide cost-effective solutions to address specific service needs while minimizing cost.
- Support Achievement: retention incentives should successfully compete for talent and reward exceptional performance.

We would add two additional attributes to the QRMC criteria:

- Precision: retention incentives should accurately meet their intended force-management objectives, including overall end-strength targets, balance across career fields, and distribution across specific assignments.
- Practicality: retention incentives should be easy for the services to implement and easy for service members to participate.

The current retention incentive system is voluntary and practical but scores much lower for flexibility and responsiveness, best value, supporting achievement, and precision. How well an auction can meet these attributes depends in part on the auction design. A simple auction that offers monetary bonuses to the service members submitting the lowest bids would score well as voluntary, flexible and responsive, and precise. It would score moderately well for best value and practicality, but would score lower for achievement (Coughlan & Gates, 2010). Incorporating personalized non-monetary incentives into the auction design would improve best value, but likely at the expense of practicality as the auction design becomes more complicated (Coughlan, Gates, & Zimmerman, 2011). Similarly, incorporating talent-related incentives would improve achievement, but again at the expense of practicality (Kelso, 2014; Williams, 2015).

Based on research to date, we conceptually envision several options to implement "market-based" retention decisions using different auction designs that provide different retention outcomes (purely monetary incentives to meet retention goals, incorporating non-monetary incentives to reduce retention costs, incorporating talent-related incentives to increase service member quality, etc.). We also conceptually understand the implications of different bidding processes (sealed bid, sequential updating, discriminatory prices, uniform prices, etc.). We have limited evidence, through experimental economics, that validates the conceptual design we feel would be most effective for a modest pilot-test proposal.

Our preference would be to conduct additional experiments to further substantiate the earlier experimental results over a broader population and for a more precise description of the auction objectives in the pilot-test proposal (e.g., meeting numerical retention goals, reducing retention costs using non-monetary incentives, incorporating talent-related incentives, etc.). Partnering with a naval community to run a pilot auction may be desirable to validate experimental results in a real-life high-stakes setting, where the cheap-talk element will be eliminated.

There are also some difficult operational details to resolve in designing a pilot-test auction (when the auction opens and closes, who is eligible to bid, whether the commitment is enforceable, etc.). Solving these issues and clarifying the operational details is not a trivial task. Yet the devil is in the details. A poorly designed pilot test risks failing to validate a good idea. These issues may be community-specific and should be addressed jointly by Naval Postgraduate School (NPS) faculty/students and Navy personnel experts in conjunction with preparing for a specific pilot test.

B. APPROACH 2: SURVEY THOSE WHO MADE DECISION ALREADY

There are two pieces of information that help to properly set a mid-career or other retention incentive:

1. How much will mid-career reenlistment rates decrease?
2. How much will a given bonus level or amount increase reenlistment rates?

Asking service members who have yet to make their reenlistment decision would obviously create an incentive to overstate how much they would respond to a change in the retirement annuity or a change in the bonus. However, there would be much less incentive to exaggerate among those who already made a reenlistment decision. Thus, we propose two extremely short surveys.

Survey among Current Mid-Career Reenlisters

To solve the first problem listed in the previous section, those who reenlisted could be surveyed with the following simple questions:

- Would you have reenlisted if the retirement annuity after 20 years of service were 40% of basic pay instead of 50% of basic pay?
- Would you have reenlisted if the SRB were one level lower?
- Would you have reenlisted if the SRB were two levels lower?

Survey among Mid-Career Exiters

To solve the second problem of how the bonus would affect the mid-career reenlistment rate, there could be a survey for those exiting at the mid-career point, with just three questions:

- Did you want to reenlist but were unable?
- Would you have stayed if the SRB were one level higher?
- Would you have stayed if the SRB were two levels higher?

For officers, the question would focus on certain bonus amounts (or multiples of monthly basic pay). For enlisted service members, the question would focus on the multiple of monthly basic pay (instead of SRBs). From these questions, we can determine what the bonus would have had to be to increase reenlistment rates.

The survey would be purposefully kept very short to improve the response rate. This would reduce any self-selection bias in who responds to the survey—for example, it may be those who want to voice frustration and may not be representative of those who actually left military service.

Currently, the Navy gives an exit survey to service members who have chosen to leave. It is a lengthy survey that appears to require at least 45 minutes to complete. Still, the response rates were higher than we expected: about 20% for enlisted and 35% for officers.

We propose that a random sample (perhaps 50%) of mid-career exiters be given this short survey instead.

For this approach to be useful, it would generally be more important to get a high response rate than it would be to get a very large number of respondents. A very large sample would not be useful if there were sample-selection bias. And, a low response rate may be indicative of sample-selection bias, as it may not be random who does or does not respond.

We do not have a scientific answer for an adequate response rate, but the accuracy of the prediction from the surveys is directly related to the response rate. This is why we believe that a survey that takes no more than two minutes would be needed. And, somehow getting this survey to be close to mandatory would help increase the response rate.

Advantages of This Approach

Using surveys as described here to estimate the retention effects from the changes in the retirement annuity and from any mid-career bonus has several advantages over the existing approach of using complicated retention models. These address some of the problems associated with the retention models mentioned earlier. The advantages of using these surveys include the following:

- They are based on current data rather than historical data. This avoids problems with inferring future compensation-retention with data from the past.
- There is no reliance on the multitude of assumptions associated with retention models. Rather, the data is based on a direct question of whether the service member would have changed their reenlistment decision with a given change. It does rely on the assumptions that service members respond truthfully and there is no bias in terms of who responds.
- Although we cannot rule them out, we do not conceive of any inherent systematic biases, as described in Chapter IV (e.g., bias from omitted variables, reverse causality, measurement error, excess supply, how to handle extensions, how to handle MOSs that are eliminated).

C. APPROACH 3: COLLECT SUBJECT-MATTER-EXPERT OPINIONS

The USS *Scorpion*, a skipjack-class nuclear submarine, sank in 1968. It came as a surprise, as family members had come to greet the *Scorpion* at its scheduled return-to-base date, but it never showed up. There are a few theories, all with supporting evidence, as to why the *Scorpion* sunk. But what is relevant here is how the *Scorpion* was eventually found. The Navy relied on a probability-weighting scheme of several subject-matter experts as they evaluated the data available. These subject-matter experts included navigators who knew the ocean terrain.

This is an example of a methodology referred to as information aggregation, where the goal is to aggregate disparate, decentralized information into a single useful answer. Prediction markets are one method for collecting subject-matter opinions that has been used in many settings. Prediction markets are markets in which participants can buy and sell shares for an outcome that will be determined in the future. This could be

- a yes/no outcome (such as whether the Navy misses its recruiting goal)—the price would reflect the probability of the event occurring
- a continuous-variable outcome (e.g., how many people will sign up for Gmail this year)

Participants owning shares reflecting the actual outcome, once the outcome is realized, receive "payments" for those winning shares. In some cases, participants may pay cash for the shares they purchase and receive cash rewards for winning shares. In other applications, payouts may involve non-cash rewards or simple bragging rights (Coughlan, Gates, & Arkes, 2011).

Major corporations have used these information-aggregation techniques, including pharmaceutical companies. Suppose there is a potential blockbuster drug, but company executives are questioning whether it will gain regulatory approval by a specific future date. The company could create a winner-takes-all market in which participants purchase "Yes" or "No" shares and receive one dollar per share if their prediction is correct. The market prices of Yes and No shares should sum to one dollar and range between zero and one dollar if the market is efficient. The Yes and No prices reflect the probability that the drug will receive regulatory approval by the proposed date. If Yes shares increase in price relative to No shares, it indicates the prospects for the drug have improved, and vice versa.

Research has already explored the feasibility of the Navy using prediction markets for personnel outcomes (Chinn & Huffman, 2009; Coughlin et al., 2011). One concern in this research is whether the military personnel-oriented prediction markets can elicit adequate participation to generate accurate predictions. Having sufficient knowledgeable participation requires a combination of an adequate informed trader population and adequate participation rates among those potential traders. While there is likely a large knowledgeable population for most military personnel-oriented issues, it is uncertain whether incentives are sufficient to draw traders into the prediction market and to keep them active over time. Would these markets be able to offer sufficient incentives, and what incentives might they use?

Another issue involves the outcomes that are good candidates for prediction market securities but are not self-defeating (endogenous). Outcomes for which policy-makers would likely intervene to affect the outcome based on the prediction market contract prices are self-defeating. For example, a market predicting the retention rate for naval aviators would be endogenous if a low retention prediction led to an increase in retention incentives before the prediction market closed. This endogeneity likely encompasses a large share of potential military manpower applications. It is important to define contract outcomes so that they avoid policy intervention (e.g., predict intermediary variables as opposed to policy-relevant final outcomes), so that they involve ranges of potential outcomes, or so that they include payoff criteria if policy interventions alter the underlying conditions.

Prediction markets are both powerful and complicated. They provide the best available opportunity to aggregate disparate, decentralized information yet they are complicated mechanisms posing complex implementation issues. There is substantial information regarding claim definition, claim structure, participation incentives, market participants, and trading mechanisms. At the same time, there are several factors regarding prediction market applications in Navy and Defense Department applications that are still poorly understood, including market participation and contract definition.

Nevertheless, the ideas of prediction markets can be adopted without formally creating a prediction market. In particular, opinions of informed experts can be collected. The trick would be to identify those experts, get them to think analytically about the questions at hand, and find a way to effectively aggregate their individual opinions. There are interesting machine learning approaches that might be able to endogenously adjust the weights on experts' predictions based on their historical record of accuracy. Whether through prediction markets or machine learning approaches to aggregating expert opinions, information aggregation provides an interesting but largely undeveloped approach to forecasting the impacts of military personnel policy changes.

X. DOES A FIXED BONUS (BY SERVICE AND ENLISTED-VS.-OFFICER) AND FIXED CAREER POINT WORK?

A. PROBLEMS WITH A FIXED BONUS (BY SERVICE AND ENLISTED-VS.-OFFICER)

Due to computational limitations discussed above, the RAND DRM model (Asch et al., 2015) estimated eight fixed bonuses, one for officers and one for enlisted for each of the four services. It is widely acknowledged that these bonuses would undoubtedly be inefficient for most MOSs. Various MOSs, skills, or communities will vary in the bonus needed to keep the necessary number of service members. This means that a fixed bonus across all communities would be too high or too low for most MOSs. The too-high bonuses would mean that the services are paying more than necessary; the too-low bonuses would mean that the services would not keep as many service members as it needed and would have to adjust the bonus upwards. Furthermore, the services may lose some high-quality service members with a too-low bonus.

B. PROBLEMS WITH A FIXED CAREER POINT

While the new law gives the services the latitude to pay the bonus anywhere in the 8–12-year range, all services currently plan to offer the bonus at the 12-year point.

A bonus at a fixed point that carries an extra required obligation is equivalent to a reenlistment. For a service member who is up for reenlistment at the time of the bonus, this makes sense. But, for another service member who just reenlisted for four years at YOS 11, the extra obligation that comes with the 12-year bonus would be coincidental with the extra obligation from the one-year-earlier reenlistment. We are not aware of any part of the policy that takes this issue into consideration.

To examine this issue, we examined the initial EAOS date, be it a reenlistment or loss, that occurred in a Service member’s 9th to 12th year of service for various ratings and skills. The problems are demonstrated in Figure 4. We compare the YOS that the first end-of-active-obligated-service (EAOS) date falls in the YOS 9–12 period for three ratings. The first point from the figure is that the EAOS date does not have a systematic point in that four-year period. This, of course, is not surprising, given (1) the various initial obligation lengths for sailors with different skills within a rating, and (2) the cases in which a service member reenlists well before his/her EAOS date. The second important point is that there are large differences across ratings in the distribution of where the EAOS date falls in that length-of-service period.

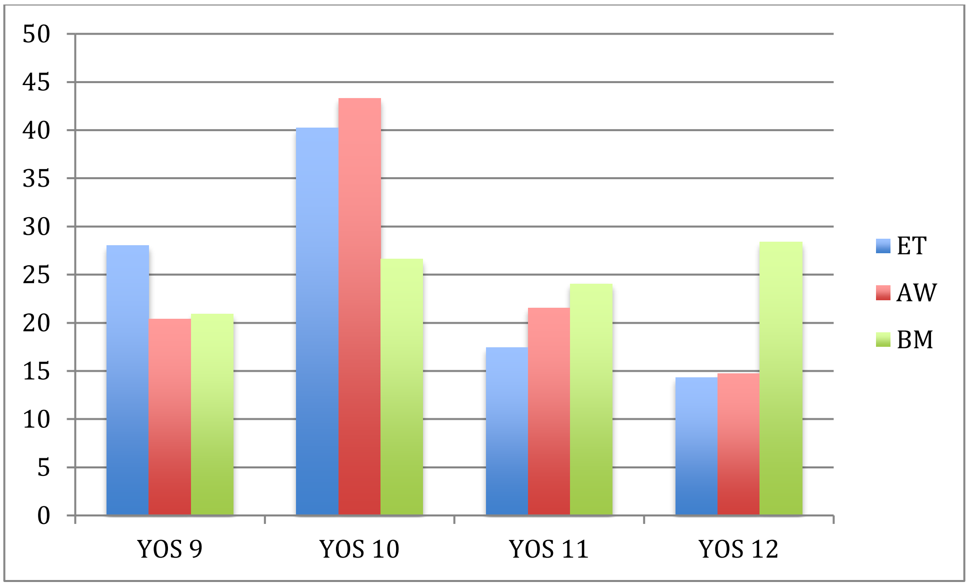

Figure 4. Variation in the Timing of ETS' across Navy Ratings

Figures 5 and 6 demonstrate differences in the distribution of those EAOS dates, within two ratings, for the two top NECs within each rating. The conclusions are the same as for differences across ratings. There is a wide variation in the EAOS date, and there are differences across skills in the distribution.

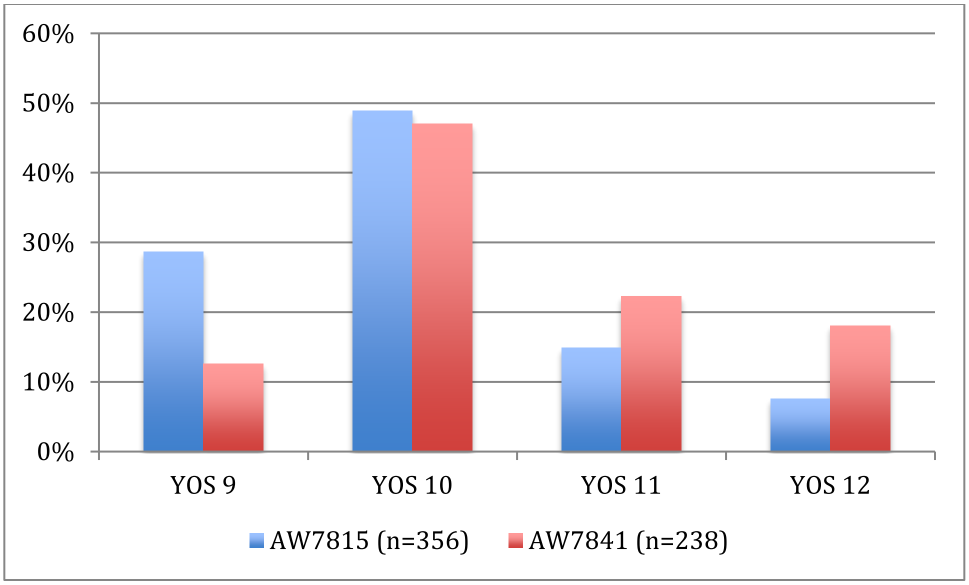

Figure 5. Variation in the Timing of ETS' across Two Skills within the AW Rating

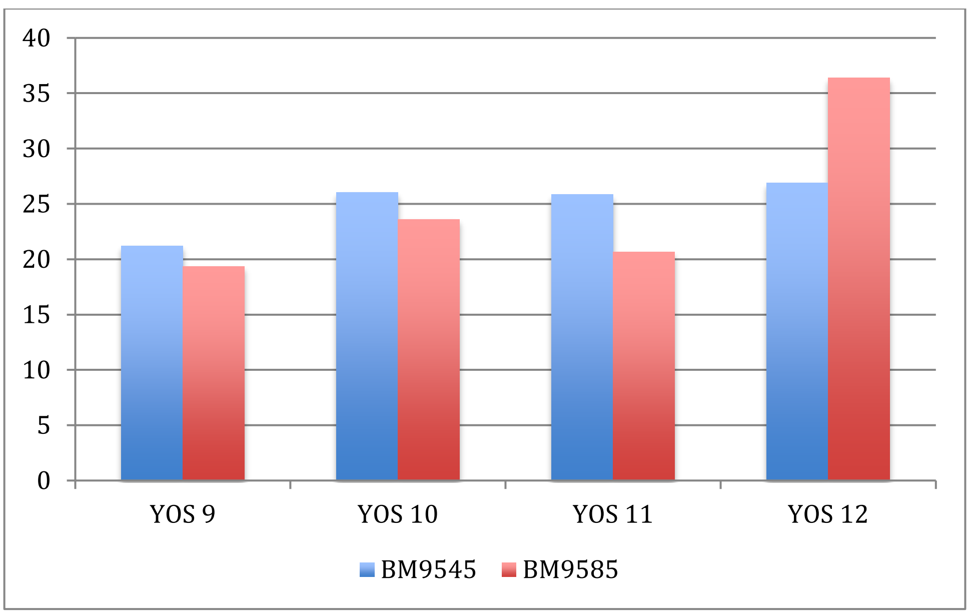

Figure 6. Variation in the Timing of ETS' across Two Skills within the BM Rating

The bottom line from this analysis is that a fixed career point for the mid-career bonus would be inefficient and wasteful since EAOS dates fall at a variety of career points in that mid-career range. If a service member reenlists at the end of YOS 11 and committed through YOS 15, another bonus requiring a four-year commitment at the end of YOS 12 would only require an extra one-year commitment. Having flexibility in when the bonus is given would allow the Navy (and other services) to achieve more bang for the buck from the mid-career bonus.

XI. CONCLUSIONS/STEPS FORWARD

In this report, we document the significant challenges in providing precise predictions for retention outcomes in the military. The non-random nature of pays and bonuses, data limitations and mismeasurement issues, non-standard reenlistment timing rules, and demand constraints on reenlistments all contribute to biases that are present in all major models to date, resulting in systematically incorrectly-estimated impacts of retention policies.

Our understanding of these biases, advances in econometric techniques, better/more accurate/more expansive data collection, and increased computing resources will lead to the development of a suite of models that will provide the Navy with more accurate and useful short-term and long-run predictions of the size and composition of its fighting force.

Our efforts moving forward will be broadly split into two tracks. In the first track, we will develop a new dynamic programming model. Similar to the RAND DRM, this model will ultimately seek to provide steady-state estimates and transition dynamics of the impacts of changes in compensation policies. In contrast to the DRM, we will build the model from the ground up to incorporate individual-specific demographic and professional data. Recent econometric innovations at the frontier of research, such as conditional choice probability (CCP), can circumvent the need to do complicated and time-consuming computations to fully solve the model and may allow us to estimate the impact of non-monetary incentives as well as differential rates of retention by officer MOS and/or quality (Hotz & Miller, 1993). Dynamic programming models will provide a long-run perspective on retention policy. For the Navy to project force size and composition decades in advance, a model that considers the lifetime decision-making process of individuals is essential.

In the second track, alternative approaches to augment models in determining the optimal bonus will be explored. As even the most ambitious and complicated dynamic programming models will continue to suffer from some of the biases we have identified in this report, more direct observation and investigation into individual behavior may offer additional clues as to how policies would affect reenlistment decisions. In particular, in the short term, when the objective is the immediate needs of the Navy to fill billets at minimal cost, observations that are more direct may provide the sharpest policy prescriptions. We believe that the most promising of these is an auction market. We hope to collaborate with a Navy community to run a pilot auction in the near future as a proof-of-concept.

Finally, we mention one important caveat for the output from our research. The eventual outcome of any analysis determining the optimal mid-career bonus will likely assume that all non-observable factors that affect retention remain constant in our analysis period and in the future when the new bonuses are set. These non-observable factors could include, among a multitude of things, the following:

- unexpected changes in civilian earnings potential,
- changes in OPTEMPO or the type of deployments, and
- changes to the Navy's command climate.

The first two are outside of the Navy's control. The changes to the Navy's command climate are something that the Navy could control. It would be overly ambitious to assume that we could predict how these factors will evolve over years and decades. The myriad of unanticipated changes to the economic, political, social, and military environments and subsequent dynamic reactions to these changes by the Navy and all other military services, government, and civilian organizations (and the subsequent feedback to the environment) means that any estimate from our models beyond the very short term must be understood in the context of a "baseline." If the command climate worsens, then the bonus would need to be higher than what models predict based on historical data. However, if the Navy improves the command climate, then lower bonuses would be needed and models would likely overstate the bonus needed. Our models will not produce estimates that will allow a "set-it-and-forget-it" retention policy. Indeed, we would be highly skeptical of any model that would claim such a capability. Instead, our models will provide guidance for what retention will look like when the status quo is maintained. Large-scale changes in the environment will require new models and analysis to account for these factors and provide Navy leadership with more accurate predictions to assist data-driven decision-making.

APPENDIX: TECHNICAL DETAILS OF THE DYNAMIC RETENTION MODEL

Possibly the first full-blown dynamic retention model (DRM) in the military economics literature is the one developed by Gotz and McCall (1984). They used dynamic programming to analyze the stay/leave decisions of Air Force officers facing diverse compensation incentives at different moments in their careers. Unfortunately, the high level of generality of this model made solving it and estimating its structural parameters considerably difficult. One reason for this difficulty is related to the so-called Bellman's curse of dimensionality, which is a usual problem that arises with the numerical solution of dynamic programming models. In particular, the curse of dimensionality is due to the discretization of continuous state and decision variables, because the computer time and space needed to solve the dynamic model increase exponentially with the number of points in the discretization (Rust, 1997, 2008). The computational power available in the beginning of the 1980s plus the complexity of the Gotz and McCall's formulation of the DRM allowed them to estimate only three parameters (i.e., the mean and standard deviation of the constant component of the taste parameter plus the standard deviation of the zero-mean transitory component of the taste parameter). The estimation of the model was thus restrictive in some dimensions. For instance, they did not compute the standard errors for the three estimated parameters; they fixed the discount rate, and did not include observable exogenous covariates.

To make the DRM more tractable, Daula and Moffit (1995) introduced a few simplifications that made the model solution and, thus, estimation a bit easier. In this manuscript, we describe a simplified version of the original DRM, which is similar to the one in Daula and Moffit (1995). We start describing the decision-maker's problem.

We assume service members are rational agents who make career choices to maximize their lifetime utility. In particular, we suppose that the individual weighs all the costs and benefits in each decision, including both the monetary and non-monetary components. In the present problem, at the end of each period of service, the military member decides whether to stay in the force for another term or to leave for the civilian sector. We further assume that returning to the military after leaving is not possible, so the leave decision is irreversible.

The monetary components that the individual considers include (1) regular military compensation (i.e., basic pay, allowances, and tax benefits), selective reenlistment bonuses, and military retirement; and (2) civilian compensation. The non-pecuniary components include the individuals' taste or preference for the military versus the civilian style of life. In this sense, some individuals prefer the type of employment and patriotism involved in the military service to any civilian activity. Therefore, there will be service members who prefer to stay in the military even when they could increase their income by choosing to work in the civilian sector. On the contrary, there will be individuals who prefer civilian jobs even if they could obtain higher income in the military. We introduce these non-monetary components into the model using "taste parameters" that reflect the monetary-equivalent of service members' preferences for the military and civilian lives.

The basic notation includes the following:

- W_t^m indicates the regular military compensation that the individual can obtain in the military in year of service t (including bonuses)
- W_t^c denotes the compensation that the service member can obtain in the civilian sector in year of service t (including military retirement)
- T represents the time horizon of the decision problem (e.g., the expected number of periods until final retirement)[7]
- $\beta = \frac{1}{1+r}$ indicates the discount factor, and r is the subjective discount rate of the service member
- ω^c denotes the taste parameter that captures the monetary equivalent of the preference for the civilian life
- ω^m denotes the taste parameter that captures the monetary equivalent of the preference for the military service
- $E_t[.]$ is the expectation operator given the information in year of service t
- ε_t^c and ε_t^m are random variables with zero mean

[7] To avoid further complexities, we assume that the service member's income after final retirement (i.e., beyond horizon T) is the same regardless of his/her stay/leave choices. Thus, we can disregard this stream of payments.

We let super-index L denote the decision to leave for the civilian sector and super-index S refer to the decision to stay in the military. Then, the service member's problem can be written as

$$V_t^L = W_\tau^c + \omega^c + \beta E_t[V_{t+1}^L] + \varepsilon_t^c = \sum_{\tau=t}^{T} \beta^{\tau-t}(W_\tau^c + \omega^c) + \varepsilon_t^c, \quad (1)$$

$$V_t^S = W_t^m + \omega^m + \beta E_t[V_{t+1}] + \varepsilon_t^m, \quad (2)$$

$$V_t = Max[V_t^L, V_t^S] \quad (3)$$

where V_t^S denotes the present value of staying in the military for another term while V_t^L indicates the present value of leaving the military to pursue a civilian career. According to this setup, the service member will remain in the military as long as the value of staying, V_t^S, exceeds the value of leaving, V_t^L. This decision problem refers to a specific service member and, thus, all variables and parameters are individual specific.

One of the main advantages of the DRM is that it yields a service member's behavior that is time consistent. By contrast, a behavior is dynamically or time inconsistent when the initial "optimal" plan of action is no longer optimal as time passes and new information arrives (e.g., new realizations of ε_t^c and ε_t^m occur).[8] In the DRM, the initial optimal plan of action remains optimal in all the future decision points.

Nevertheless, the version of the DRM described above has some disadvantages. The latter are given by the still strong assumptions made about the agent's behavior. For example, the service member is assumed to forecast the future income streams (i.e., W_t^m and W_t^c) as well as the time horizon T with certainty. We also suppose the member knows the distribution functions for the errors (i.e., ε_t^c and ε_t^m) and their current values, and that the discount factor β is not time-varying. In addition, we assume the taste parameters ω^c and ω^m are constant over the life of the service member and that the individuals have linear utility functions (i.e., in Equations 1 and 2 we are discounting cash flows instead of "non-linear utility flows"). While all these assumptions could be relaxed, the added complexity would make the model solution and estimation increasingly difficult. Finally, most of the disadvantages just described are also shortcomings of the other retention models (i.e., TCOL, ACOL, and ACOL 2).

[8] In other words, the original policy does not satisfy Bellman's Principle of Optimality.

THIS PAGE INTENTIONALLY LEFT BLANK

REFERENCES

Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics, Vol. 3*. Amsterdam, The Netherlands: Elsevier.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics*. Princeton, NJ: Princeton University Press.

Arkes, J. (2016). Empirical biases and some remedies in estimating the effects of selective reenlistment bonuses on reenlistment rates. *Defence and Peace Economics*, 29(5), 475-502.
.

Arkes, J. (2017). *Regression analysis: A practical introduction*. Unpublished manuscript.

Asch, B. J., Heaton, P., Hosek, J., Martorell, F., Simon, C., & Warner, J. T. (2010). *Cash incentives and military enlistment, attrition, and reenlistment*. Santa Monica, CA: RAND.

Asch, B. J., Johnson, R. W., & Warner, J. T. (1998). *Reforming the military retirement system*. Santa Monica, CA: RAND.

Asch, B. J., Mattock, M. G., & Hosek, J. (2013). *A new tool for assessing workforce management policies over time: Extending the dynamic retention model*. Santa Monica, CA: RAND.

Asch, B. J., Mattock, M. G., & Hosek, J. (2015). *Reforming military retirement: Analysis in support of the Military Compensation and Retirement Modernization Commission*. Santa Monica, CA: RAND. Retrieved from https://www.rand.org/pubs/research_reports/RR1022.htm

Asch, B. J., Mattock, M. G., & Hosek, J. (2017). *The blended retirement system: Retention effects and continuation pay cost estimates for the armed services*. Santa Monica, CA: RAND.

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.

Black, M., Moffitt, R., & Warner, J. T. (1990). The dynamics of job separation: The case of federal employees. *Journal of Applied Econometrics, 5*(3), 245–262.

Chinn, M. A., & Huffman, L. A. (2009). *Prediction markets: A review with an experimentally-based recommendation for Navy force-shaping application* (MBA report). Monterey, CA: Naval Postgraduate School.

Coughlan, P. J., & Gates, W. R. (2010). Auctions as force-shaping and force-management tools. In J. Parco, D. Levy, & F. Randy (Eds.), *Attitudes aren't free: Thinking deeply about diversity in the US Armed Forces*. Maxwell Air Force Base, AL: Air University Press.

Coughlan, P. J., Gates, W. R., & Arkes, J. (2011, February 3). *Innovations in defense acquisition: Asymmetric information, mechanism design and prediction markets* (NPS-AM-11-06). Monterey, CA: Naval Postgraduate School, Acquisition Research Program.

Coughlan, P. J., Gates, W. R., & Zimmerman, B. M. (2011). *The Combinatorial Retention Auction Mechanism (CRAM): Integrating monetary and non-monetary reenlistment incentives* [Technical report] (NPS-GSBPP-1-013). Monterey, CA: Naval Postgraduate School.

Daula, T., & Moffitt, R. (1995). Estimating dynamic models of quit behavior: The case of military reenlistment. *Journal of Labor Economics*, *13*(3), 499–523.

Enns, J. H. (1975). *Effect of the variable reenlistment bonus on reenlistment rates: Empirical results for FY 1971* (Vol. 1502, No. ARPA). Santa Monica, CA: RAND.

Gibbons, C., & Suarez Serrato, J. C. (2011). *Broken or fixed effects?* [Working paper].

Gotz, G. A., & McCall, J. J. (1984). *A dynamic retention model for Air Force officers: Theory and estimates* (R-3028-AF). Santa Monica, CA: RAND.

Hansen, M. L., & Wenger, J. W. (2002, November). *Why do pay elasticity estimates differ?* (CRM D0005644.A2/Final). Alexandria, VA: Center for Naval Analyses.

Hansen, M. L., & Wenger, J. W. (2005). Is the pay responsiveness of enlisted personnel decreasing? *Defence and Peace Economics*, *16*, 29–43.

Hattiangadi, A. U., Ackerman, D., Kimble, T. H., & Quester, A. O. (2004 , May). *Cost-benefit analysis of lump sum bonuses for Zone A, Zone B, and Zone C reenlistments: Final report* (CRM D0009652.A4/1Rev). Alexandria, VA: Center for Naval Analyses.

Hogan, P., & Black, M. (1991). Reenlistment models: A methodological review. In C. L. Gilroy, D. K. Horne, & D. A. Smith (Eds.), *Military compensation and personnel retention: Models and evidence*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hosek, J. R., & Martorell, F. (2009). *How have deployments during the war on terrorism affected reenlistment?* (MG-873-OSD). Santa Monica, CA: RAND. Retrieved from http://www.rand.org/pubs/monographs/MG873/

Hosek, J. R., Mattock, M. G., Fair, C. C., Kavanagh, J., Sharp, J., & Totten, M. (2004). *Attracting the beast: How the military competes for information technology personnel*. Santa Monica, CA: RAND.

Hosek, J. R., Peterson, C. E. (1985). *Reenlistment bonuses and retention behavior* (R-3199-ML). Santa Monica, CA: RAND.

Hotz, V. J., & Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies, 60*(3), 497–529.

Kelso, E. W. (2014). *Improving the efficiency of aviation retention bonuses through the use of market mechanisms* (MBA Joint Applied Project). Monterey, CA: Naval Postgraduate School.

Mackin, P. C. (1996). *Reestimation of ACOL coefficients for the CAPM model: Final report* (Project No. HQ0038-4108-0012). Falls Church, VA: SAG.

Mattock, M., & Arkes, J. (2007). *The dynamic retention model for Air Force officers: New estimates and policy simulations of the Aviator Continuation Pay Program* (RAND-TR-470-AF). Santa Monica, CA: RAND.

Rust, J. (1997). Using randomization to break the curse of dimensionality. *Econometrica, 65*(3), 487–516.

Rust, J. (2008). Dynamic programming. In *The new Palgrave dictionary of economics* (2nd ed.) London, England: Palgrave Macmillan.

Under Secretary of Defense for Personnel and Readiness. (2008). *10th quadrennial review of military compensation, Vol. 1*. Washington, DC: Author.

Warner, J. T. (1979). *Alternative military retirement systems: Their effects on enlisted retention*. Alexandria, VA: Center for Naval Analyses.

Warner, J. T., & Goldberg, M. (1984). The influence of non-pecuniary factors and labor supply: The case of Navy enlisted personnel. *Review of Economics and Statistics*, *66*(l), 26–35.

Williams, B. M. (2015). *Evaluation of aviation career pay incentives among the naval aviation enterprise utilizing auction mechanisms* (Master's thesis). Monterey, CA: Naval Postgraduate School.

Wooldridge, J. M. (2005). Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models. *Review of Economics and Statistics*, *87*(2), 385–390.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Research Sponsored Programs Office, Code 41
   Naval Postgraduate School
   Monterey, CA 93943