



# NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

**DETERMINANTS OF NAVY PERFORMANCE EVALUATION**

**RECORDS**

by

Sae Young Ahn and Latika Hartmann

October 2023

**Approved for public release. Distribution is unlimited.**

Prepared for: N1 – Manpower, Personnel, Training and Education  
This research is supported by funding from the Naval Postgraduate School Naval  
Research Program

THIS PAGE INTENTIONALLY LEFT BLANK

## REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

<b>1. REPORT DATE</b> October 3, 2023	<b>2. REPORT TYPE</b> Technical Report	<b>3. DATES COVERED</b>	
		<b>START DATE</b> 10/23/2022	<b>END DATE</b> 10/21/2023
<b>4. TITLE AND SUBTITLE</b> Determinants of Navy Performance Evaluation Records			
<b>5a. CONTRACT NUMBER</b>	<b>5b. GRANT NUMBER</b>	<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>5d. PROJECT NUMBER</b> NPS-23-N102-A	<b>5e. TASK NUMBER</b>	<b>5f. WORK UNIT NUMBER</b>	
<b>6. AUTHOR(S)</b> Sae Young Ahn and Latika Hartmann			
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Naval Postgraduate School 1 University Circle Monterey, CA 93943-5000			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>  NPS-DDM-23-005
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> N1 – Manpower, Personnel, Training & Education		<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> N1	<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b> NPS-23-N102-A
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> Approved for Public Release. Distribution is Unlimited.			
<b>13. SUPPLEMENTARY NOTES</b> The views expressed in this document are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
<b>14. ABSTRACT</b> Using detailed data on evaluation and promotion outcomes of active-duty sailors and officers from 2014 to 2022, we find that earning a recommendation of Early Promote (EP) and receiving individual trait average scores above the reporting senior cumulative average (RSCA) are most predictive of promotion across ranks and communities. We find mixed results for socio-demographic and evaluation characteristics assumed to have a negative impact on promotion. Being non-white or a "1 of 1" evaluation (non-competitive summary group) is negatively associated with promotion in some rank and community combinations. Being female or married is positively associated with promotion in some rank and community combinations. However, these results are inconsistent across all ranks and communities, and we fail to find consistent patterns in the data that point to racial or gender bias substantively disadvantaging minorities or women.			
<b>15. SUBJECT TERMS</b>			
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>
<b>a. REPORT</b> Unclassified	<b>b. ABSTRACT</b> Unclassified	<b>c. THIS PAGE</b> Unclassified	<b>UU</b>
			<b>18. NUMBER OF PAGES</b>  66
<b>19a. NAME OF RESPONSIBLE PERSON</b> Dr. Latika Hartmann			<b>19b. PHONE NUMBER (Include area code)</b> (310) 804-1916

THIS PAGE INTENTIONALLY LEFT BLANK

**NAVAL POSTGRADUATE SCHOOL  
Monterey, California 93943-5000**

Ann E. Rondeau  
President

Scott Gartner  
Provost

The report entitled “Determinants of Navy Performance Evaluation Records” was prepared for N1 – Manpower, Personnel, Training and Education and funded by the Naval Postgraduate School Naval Research Program.

**Further distribution of all or part of this report is authorized.**

**This report was prepared by:**

AHN.SAE  
YOUNG.1539  
252565

Digitally signed by  
AHN.SAE  
YOUNG.1539252565  
Date: 2023.10.03  
13:41:19 -07'00'

Sae Young Ahn  
Assistant Professor

---

HARTMANN.L  
ATIKA.150170  
0025

Digitally signed by  
HARTMANN.LATIKA.150  
1700025  
Date: 2023.10.03  
13:08:03 -07'00'

Latika Hartmann  
Associate Professor

---

**Reviewed by:**

Raymond  
Jones

Digitally signed by Raymond  
Jones  
Date: 2023.10.04 16:22:26  
-07'00'

Ray Jones, Chairman  
Department of Defense Management

**Released by:**

---

Kevin B. Smith  
Vice Provost for Research

THIS PAGE INTENTIONALLY LEFT BLANK

# TABLE OF CONTENTS

<b>I. INTRODUCTION .....</b>	<b>1</b>
<b>A. BACKGROUND .....</b>	<b>1</b>
<b>B. FINDINGS .....</b>	<b>3</b>
<b>II. DATA AND SUMMARY STATISTICS .....</b>	<b>6</b>
<b>A. DATA ORGANIZATION .....</b>	<b>6</b>
<b>B. SUMMARY STATISTICS.....</b>	<b>7</b>
<b>III. MODEL .....</b>	<b>12</b>
<b>IV. RESULTS .....</b>	<b>18</b>
<b>A. O-3 TO O-4 PROMOTIONS .....</b>	<b>18</b>
<b>B. O-4 TO O-5 PROMOTIONS .....</b>	<b>21</b>
<b>C. O-5 TO O-6 PROMOTIONS .....</b>	<b>23</b>
<b>D. DISCUSSION OF OFFICER PROMOTION RESULTS .....</b>	<b>24</b>
<b>E. ENLISTED RANKS.....</b>	<b>26</b>
<b>V. CONCLUSION AND RECOMMENDATIONS.....</b>	<b>29</b>
<b>TABLES.....</b>	<b>32</b>
<b>APPENDIX TABLES.....</b>	<b>48</b>
<b>LIST OF REFERENCES.....</b>	<b>53</b>
<b>INITIAL DISTRIBUTION LIST .....</b>	<b>56</b>

THIS PAGE INTENTIONALLY LEFT BLANK

# I. INTRODUCTION

## A. BACKGROUND

Identifying and promoting the “best and brightest” is essential to the productivity of any organization. For the U.S. Navy, these aims are even more critical because the Navy is a closed personnel system with virtually no entry of senior personnel from outside the organization. Given these circumstances, it is essential for the Navy to find and nurture talent. Thus, performance evaluation, retention of productive personnel, and promotion of enlisted sailors and commissioned officers into billets in the hierarchy that allows them to succeed within the U.S. Navy is both an incredibly critical but daunting task.

As the Navy moves forward with transforming the Navy Performance Evaluation system, this report serves as a follow-on study to Ahn and Hartmann (2022) where we compared the Navy evaluation instruments (i.e., fitness reports [FITREPs] and evaluations [EVALs]) to the other services, solicited focus group feedback from Navy selection board members, and fielded a survey to the larger Navy enlisted and officer populations eliciting feedback on proposed changes to the system. We build on that work in this report and use detailed individual records on promotion and performance evaluation to provide a rigorous examination of how the current promotion system operates. Combining individual data from 2014 to 2022, we examine the key predictors of officer promotions and enlisted sailor time to promotion from 2014 to 2022. This research supports current Performance Evaluation Transformation efforts being led by Navy Personnel Command and Task Force One Navy #2 on Talent Management.

To conduct the data analysis, we merged and operationalized multiple sources of data from Navy Personnel Command (NAVPERS), Defense Manpower Data Center (DMDC), and MyNavyHR to construct a longitudinal record of active-duty Navy enlisted sailor and officer promotion candidates of ranks E-5 and O-3 to O-5 from 2014 to 2022, with their detailed promotion information, evaluation and career characteristics, and demographics. With these data, we conducted regression analysis to identify the key drivers of promotions from factors related to the most recent evaluation and past evaluations, career, and demographic characteristics. For officers, apart from differences by rank, we also explored differences between the unrestricted line (URL) and restricted

line plus staff corps (RL-Staff) as a single aggregate group. Within the URL, we looked at the larger communities of Aviation and Surface Warfare, while within the RL-Staff we looked at the large, combined communities of Medical and Judge Advocate General's Corps (MED-JAG) as a single group. For enlisted sailors, we separated sailors by ratings to look for potential differences by occupation group.

Unlike the extensive literature on military retention and attrition (e.g., Marrone, 2020), the literature on military performance evaluation systems is small.<sup>1</sup> A few studies look at the process surrounding military performance evaluation, in particular the Navy and Air Force process, and suggest changes, in some cases based on civilian evaluation systems (Wolfgeher, 2009; Ellison, 2014; Small, 2020). Others use evaluation records to identify particular problems such as misevaluation of performance in the Army (Evans, 2018; Evans & Bae, 2019), inflated evaluations in the Air Force (Baker, 2017), and bias in the USMC (Dunst, 2018). In particular, the USMC studies find different USMC occupations value different qualities in their evaluation suggesting less of a uniform standard for promotion (Clemens et al., 2012). Among Enlisted Marines, Larger (2017) finds proficiency and conduct marks are the most predictive of composite evaluation scores and subsequent promotion. Dunst (2018) finds white reporting seniors rate white Marines more favorably, while non-white reporting seniors rate non-white Marines relatively low, although the gaps shrink after accounting for education and combat experience pointing to the importance of controlling for all observable differences in characteristics that may be correlated with gender and performance.

Apart from Ahn et al. (2020), there has been no comprehensive study of the Navy's officer promotion systems that encompasses all the communities and multiple ranks, and especially analyzes how the promotion board operates. In their study, Ahn et al. (2020) find that removal of zone designation information increases promotion chances of below and above zone candidates, implying that promotion boards have relied on the judgement of past boards to reject above zone candidates and assumed that future boards will promote the high performing below zone candidates they failed to promote.

The small number of Navy studies is partly due to data availability where most studies evaluate promotion based on a change in rank reported in the DMDC records

---

<sup>1</sup> Marrone (2020) offers an excellent overview of this literature.

without accounting for the board process. Unlike these studies, we construct an individual-board level dataset that includes multiple observations for individuals that were re-evaluated by subsequent selection boards if they were not promoted at their first in zone consideration matched to their most recent evaluation characteristics and other individual attributes.

## **B. FINDINGS**

Our principal model is a linear probability model (LPM). The outcome variable is an indicator variable which equals 1 if the service member was promoted, and 0 if not. The explanatory variables include the service member's most recent and past performance evaluation information, career characteristics such as award information and additional qualification designations (AQDs), and demographics. We also include fixed effects for age, community/rating, board year, and board look. These fixed effects help to account for some potentially relevant yet unobserved information about the service member or promotion board. For example, there may be large differences in promotion rates across communities or zone designation. Without these fixed effects, the baseline promotion rate would be estimated to be identical across all communities/ratings, and the model would predict that an officer has the same chance of promotion regardless of their zone designation. Including fixed effects captures these differences.

Combining the detailed individual records with an LPM of promotion (indicator variable) as the main outcome, we find difference between average trait scores and the reporting senior's cumulative average (RSCA) along with an Early Promote (EP) promotion recommendation on the most recent FITREP/EVAL are highly predictive of promotions. These two predictors are most likely to lead to promotion across ranks and communities. More importantly, the models suggest top performers under reporting seniors with high RSCAs (i.e., "easy evaluators") are penalized compared to those under reporting seniors with low RSCAs (i.e., "harsh evaluators") because it is more difficult for "easy evaluators" to signal top performers.

We also find substantive differences in promotion and drivers of promotions across communities and ranks. Despite the common factors above, promotion rates differ across ranks by as much as 23 percentage points (69% for O-3 to O-4 and 46% for O-5 to O-6) and 20 percentage points across communities (84% for Surface Warfare and 64%

for Aviation at O-3 to O-4). The impact of an EP (compared to the baseline of Promote) can be as small as 18 percentage points (Aviation O-3) and as large as 27 percentage points (SWO O-3).

Women and minorities are under-represented. The under-representation becomes exacerbated at higher ranks. For example, while women make up 18% of promotion candidates from O-3 to O-4, they comprise 13% of O-5 to O-6 candidates. Similarly, while non-white officers make up 27% of O-3 to O-4 promotion candidates, they comprise 22% of O-5 to O-6 candidates. We find mixed results for socio-demographic characteristics and some evaluations assumed to negatively impact promotion. Being non-white and a “1 of 1” evaluation is sometimes negatively associated with promotion. Being female or married is sometimes positively associated with promotion. However, these results are inconsistent across ranks and communities.

We fail to find “smoking guns” of bias by looking at promotions of high and low performers. We define high performers as having all EPs and low performers as never having an EP across all observable evaluations. In these groups, race and gender are insignificant drivers of promotions. Had we found such an impact, this would imply that even in clear-cut cases, demographics could predictably influence promotions. Then, race or gender bias would be the most probable cause. This is not what we find, suggesting other omitted factors correlated with race and gender may explain some of the demographic results.

Promotions become less predictable at higher ranks. While the impacts described qualitatively hold for higher ranks, their magnitudes and statistical significance fluctuate. This finding, along with the mixed results for socio-demographic characteristics strongly imply again that we are missing important factors that drive promotions. We suspect that information on the officer summary report, written comments in evaluations, ability and experience of the selection board briefer, and composition of the selection board may affect promotion and are uncaptured in the statistical analysis.

Based on our findings, we recommend the Navy mandate a ceiling for RSCA so that high performers working under reporting seniors (RS) with high RSCAs are not disadvantaged. Average trait score above RSCA is one of the most salient predictors of promotion. However, a high RSCA restricts the ability of high performers to “stand out.”

This accords with our focus group findings in Ahn and Hartmann (2022) where respondents recommended better training of reporting seniors to curb inflated evaluations and instill more honesty in the evaluation process. Mandating an RSCA ceiling would be one way to curb inflated evaluations.

We also recommend the Navy publicize promotion rates and factors positively associated with promotion to all participants in the promotion process, including the sailors/officers being evaluated, RSs, briefers, and board members. In particular, differences across ranks and communities should be clarified, and factors that play a large (and not-so-large) role in promotion should be explained so that all participants have the most complete and accurate information. We offer suggestions for future research of selection boards and data collection efforts in the last chapter.

The rest of the report is organized as follows. Chapter II describes the data organization and summary statistics. Chapter III describes the regression model and Chapter IV summarizes the main findings by rank and community. Chapter V concludes with our recommendations for reforms and data collection efforts.

## II. DATA AND SUMMARY STATISTICS

### A. DATA ORGANIZATION

The data for this project comes from two sources: Navy Personnel Command (NAVPERS) and Defense Manpower Data Center (DMDC).<sup>2</sup> Broadly, the NAVPERS data contains performance evaluations (FITREPs/EVALs), promotion board outcomes, and other professional data, such as lists of AQDs and awards held by service members. The DMDC data contains socio-demographic characteristics of all active-duty personnel.

The board outcome data forms the backbone of our dataset. The promotion outcome is extracted from the NAVPERS “inventory” data, which contains approximately 57 million monthly snapshots of all active duty service members from 2010 to 2022. Each snapshot contains some professional information, such as rank/grade, pay entry base date (PBED), lineal number, and a promotion status. The promotion status is split into three categories for officers: promoted/failed to promote/not considered for promotion. Enlisted sailors’ status never changes from “not considered,” even when they are promoted. Approximately 97% of the monthly snapshots have a status of “not considered.”

The promotion status changes to “promoted” or “failed to be promoted” the first time an officer is considered by a board. The status remains “promoted” for several months to one year, until the status is reset to “not considered,” until the next time an officer is considered for promotion to the next rank, several years later. We speculate that this may be the period between promotion and “putting on” their new rank. Thus, it is straightforward to identify the year and month of a successful promotion.

When an officer fails to promote, status is set to “fail,” and it does not get reset until a successful promotion. We often observe officers with 12 months to 60 months of continuous “fail” status in the data. Thus, while it is easy to identify the year and month of the first time an officer fails to promote, subsequent failures are difficult to distinguish from continuation status. We assume that any “failure” status beyond 36 months is

---

<sup>2</sup> These data were obtained from NAVPERS with DMDC taking the lead in merging the performance and promotion data to socio-demographic data housed at DMDC. The complete data was then shared by DMDC with the data with the PIs.

continuation status and do not use observations beyond this point to extract board outcomes. We also drop from our data a small number of instances of rank movements downward (demotions) or increases by more than one increment (data entry error or extraordinary circumstance). We use the rank/grade data in the monthly snapshots to calculate time to promote for both enlisted sailors and commissioned officers. We then retain only the year and month snapshot when a board outcome or a change in rank (for enlisted sailors) occurs. The number of observations decreases from 57 million to 1.7 million.

Each promotion outcome observation is then merged with data from the service member's last five performance evaluations. The data for each performance evaluation arises from the evaluation data set, which contains approximately 2.9 million records. Each observation in the evaluation dataset corresponds to a FITREP/EVAL received by a service member between 2014 and 2022. It contains the beginning and ending dates of the evaluation, the average of the seven trait scores, the identity of the RS, the RSCA and the number of evaluations completed by the RS, and the promotion recommendation (EP/MP/P, etc.). To each FITREP/EVAL observation, we attach socio-demographic and professional information of the ratee, including race, gender, marital status, dependent count, AFQT (for enlisted sailors), education level, community/rating, number of AQDs earned, and award information. Finally, we drop promotion outcomes in non-competitive ranks (E-2 to E-5 and O-2 to O-3) and for senior leaders (E-7 to E-9 and O-7 to O-9).

## **B. SUMMARY STATISTICS**

We begin by reviewing summary statistics on promotion, evaluation and career characteristics, and demographics by officer rank, military occupation specialty code (MOS), and demographics. Although we include some enlisted sailor analysis, most of this study centers on junior and mid-grade officer promotion outcomes. Within MOS, we disaggregate the data for the URL and RL along with Staff Corps as a single group. We refer to the latter group as RL-Staff in the text below. We also study the bigger communities separately, namely Aviation and Surface Warfare (SWO), and the Medical and Judge Advocate General's Corps (MED-JAG). As described in Chapter III below, the unit of observation is the individual-board for selection boards conducted between 2014 and 2022 for ranks O-3 to O-4, O-4 to O-5, and O-5 to O-6. We include multiple

observations per individual to capture personnel that may not promote in zone and are reconsidered for promotion when they are above zone.

As seen in Table 1A, 69% of officers being considered for O-4 (i.e., Lieutenant Commander), promote albeit with differences by MOS and demographics. URLs are more likely to promote at 72% compared to the RL-Staff at 67%. We also find significant differences within the URL as 84% of SWOs promote compared to 64% of officers in Aviation. Among the RL-Staff, MED-JAG officers promote at higher rates of 72% compared to the average for RL-Staff officers (67%).

Apart from Aviation, the evaluation characteristics are similar across the communities with average trait evaluations on the most recent FITREP ranging from 4.3 for Aviation officers compared to 4.1 for MED-JAG officers with similar differences between the individual trait averages and the reporting senior's cumulative average (RSCA). On average, 65% of candidates receive an Early Promote (EP) on their most recent evaluation and 24% receive a Must Promote (MP). O-3 candidates up for promotion have 4.8 awards earned over the last five years on average with URLs receiving more awards compared to their RL-Staff counterparts (5.3 versus 4.4). Only 22% of Aviation officers have a graduate degree compared to 59% of SWOs and 61% of MED-JAG. In the case of Aviation, flight hour requirements probably impinge on the ability of officers to pursue graduate education early in their career while for MED-JAG officers a professional degree is often required. Although prior enlisted officers account for 12% of O-3 candidates, they are overrepresented in the RL-Staff at 16% compared to only 6% in the URL.

In terms of demographics, women account for 18% of O-3 candidates ranging from a low of 9% of Aviation officers compared to 39% of MED-JAG officers. Such demographic differences by MOS are starker for African Americans with only 3% of African American officers in Aviation compared to 11% in Surface Warfare and 12% in RL-Staff. We explore these differences further in Table 1B that summarizes promotion and evaluation characteristics by gender and race. Women promote at slightly higher rates compared to men (71% versus 69%), while African Americans promote at significantly lower rates compared to white candidates (61% versus 71%). Promotion rates for Non-Black minorities lie between the two groups at 65%.

In terms of evaluation, the gender and racial differences in EP, MP and trait average are small. Rather, the promotion differences may be related to differences between an individual's trait average and the RSCA as this delta, a proxy of a reporting senior's relative assessment of an individual, is smaller for African Americans than white candidates (0.17 versus 0.21). Furthermore, African American candidates are more likely to be prior enlisted (22% versus 10% for white candidates). Both these factors may be negatively correlated with promotion to the next rank and may account for part of the racial gap in promotion. While minorities are more likely to have graduate education, the gender difference is larger than the racial difference: 48% of women candidates have a graduate education compared to 38% of men, which is likely related to the over representation of women in the RL-Staff. Female candidates also have fewer dependents than their male counterparts (0.8 versus 1.4).

Table 2A summarizes the same characteristics for O-4 candidates up for O-5 (i.e., Commander [CDR]). Promotion rates go down as candidates move up in rank with 58% of O-4s promoting compared to 69% of O-3s. Unlike O-3 candidates, differences in promotion within the URL and RL-Staff communities shrink with both Aviation and SWO candidates promoting at similar rates of around 60%. Yet, the difference between the URL and RL-Staff persists with 62% of URLs promoting compared to 55% of the RL-Staff. We also observe differences in evaluation characteristics by MOS with 66% of SWO and MED-JAG candidates receiving EPs compared to 72% for other communities. O-4 candidates are also more likely to be part of non-competitive summary groups (1 of 1) compared to O-3 candidates (37% of O-4s are 1 of 1 compared to 33% of O-3s). O-4 candidates earn more awards and qualifications than O-3s. Graduate education and prior enlisted status are other areas of difference with 69% of O-4 candidates having a graduate education compared to only 40% of O-3s. Additionally, only 1% of O-4 candidates are prior enlisted, though that may be related to age because prior enlisted officers are significantly older when they move to the officer rank.

The Navy loses more women as officers transition from O-3 to O-4 with women accounting for 15% of O-4 candidates compared to 18% of O-3 candidates with larger drop-offs in the URL compared to the RL-Staff community. The share of Non-Black minorities also drops among O-4 candidates to 15% compared to 18% for O-3 candidates,

while the African American share remains the same at 9%. Looking at promotion differences by race and gender in Table 2B, women promote at similar rates to men, while the racial gap in promotion between African Americans and white candidates, and Non-Black minorities and white candidates persists with 49% of African American candidates promoting to O-5, 53% of Non-Black minorities promoting and 60% of white candidates. Unlike O-3 candidates where many evaluation characteristics were similar by race, Table 2B shows racial differences in evaluation characteristics with a larger share of white O-4 candidates receiving a higher trait average and an EP on their most recent FITREP compared to African American and Non-Black minority candidates. White O-4 candidates also earn more awards and qualifications than their non-white counterparts as do male candidates compared to females. Finally, while most O-4 candidates are married (86%) with an average of 1.8 dependents, there is a significant gender gap with 89% of male O-4 candidates being married with 1.9 dependents compared to 69% of females being married with 1.3 dependents on average.

Tables 3A and 3B summarize the patterns for O-5 candidates under consideration for promotion to O-6 (i.e., Captain). Unlike O-3 and O-4 candidates, the pool of O-5 candidates is smaller at around 4,900 individuals considered by selection boards between 2014 and 2022. Fewer than 50% of these candidates promote to O-6. The drop is especially sharp for Aviation where only 36% of candidates promote to O-6 compared to 60% of O-4 candidates. Compared to O-4 promotion, the decrease for SWOs and the RL-Staff is smaller with 53% of SWOs and 46% of officers in the RL-Staff promoting to O-6 (compared to 59% and 55% of O-4 candidates, respectively). Differences in promotion by MOS appear to be related with the evaluation characteristics of candidates with O-5 candidates in Aviation receiving lower trait averages and a relatively smaller share of EPs compared to their SWO counterparts.

Graduate degrees are commonplace among O-5 candidates with more than 3 out of 4 candidates having them, while the demographics of O-5 candidates are similar to O-4, albeit O-5 candidates are five years older on average (46 years) with more dependents (2 on average). Unlike O-3 and O-4 candidates, the African American – white gap in promotion disappears for O-5 candidates with African American candidates promoting at slightly higher rates (48%) compared to whites (46%). In contrast, the Non-Black

minority and white promotion gap persists though it shrinks from 7 percentage points among O-4 candidates to 3 percentage points for O-5 candidates. Most evaluation characteristics are similar across demographics. While women earn a relatively lower share of EPs, at 68% compared to men at 72%, their delta between the individual trait average and the RSCA is larger than for men (0.11 compared to 0.08). Gender differences in marriage and dependents persist among O-5 candidates.

In conclusion, this summary captures the main promotion and evaluation patterns for O-3 to O-5 candidates between 2014 and 2022 highlighting patterns that are well known (racial gaps in promotion) and others that are, perhaps, less well known. As we would expect, promotion rates decline with rank with O-3 candidates promoting at higher rates (69%) than O-5 candidates (46%). Differences in promotion by MOS are not uniform by rank with SWOs exhibiting relatively higher promotion rates among O-3 and O-5 candidates than the larger URL community while Aviation promotion rates lag behind the larger URL community among O-3 and O-5 candidates. In contrast, the RL-Staff community promotes at slightly lower rates among O-3 and O-4 candidates, while promoting at comparable rates to the URL community among O-5 candidates. Women promote at similar rates to men across the three ranks, although there are differences by gender in the share of women across MOS (the share of women in the RL-Staff is higher at each rank than the share of women in the URL), in marital status and in the number of dependents. While there are significant racial differences in promotion among O-3 and O-4 candidates, these gaps disappear for African Americans among O-5 candidates and shrink for Non-Black minority O-5 candidates compared to their white counterparts. Building on this descriptive view, the next section outlines a statistical model to assess the correlations between promotion, evaluation characteristics, career characteristics and demographics of officers in a single framework.

### III. MODEL

Our analysis focuses on promotions from O-3 to O-4, O-4 to O-5 and O-5 to O-6. We do not analyze promotions to O-2 and O-3 because the evaluation data shared by NAVPERS are not comprehensive for these ranks and promotions at lower ranks are a function of time in grade. We also do not consider promotions after O-6 because of small samples.

To identify the predictors of officer promotion, we model promotions separately by rank and evaluate the role of the most recent evaluation characteristics as reported on the FITREP, an officer’s career characteristics, demographics, and past evaluation characteristics controlling for the year of the selection board, the two-digit MOS (first two digits) of the officer, board look, and indicators for past missing evaluations. By board look, we are attempting to capture Navy officers’ zone designation. The data received from NAVPERS does not explicitly indicate whether the candidate is below/in/above zone. As such, we proxy for zone by counting the number of times an officer is assumed to be evaluated by a promotion board. Our model draws on extended discussions with Navy subject matter experts on selection boards and our review of Navy policy guidance related to selection and evaluation.

Specifically, we estimate a linear probability model of the following form:

$$Promote_{it} = x'_{ijt}\beta + w'_{it}\gamma + z'_{i(t-1)}\theta + \varepsilon_{it}$$

Here,  $Promote_{it}$ , is an indicator variable for whether officer  $i$  promoted in *board year*  $t \in \{2014, 2022\}$ . In the model,  $x_{ijt}$  represents a vector of the most recent performance evaluation data for officer  $i$  evaluated by RS  $j$ , including the promotion recommendation, the difference between their average trait score and RSCA, and whether the candidate is in a competitive summary group as we explain below. Furthermore,  $\omega_{it}$  represents a vector of socio-demographic and career characteristics for officer  $i$ . Note that these characteristics can change with time (for example, marital status), so it is subscripted with  $t$ . Finally,  $z_{i(t-1)}$  represents a vector of past evaluation and career characteristics, including the average number of AQDs, awards, etc. In the model,  $\beta$ ,  $\gamma$ , and  $\theta$  represent the impact of each of the variables on the likelihood of promotion.

If an officer promotes in their first evaluation by a selection board, we do not observe them again (until the next rank promotion outcome). If they do not promote, we

observe them again up to their fourth evaluation. At every evaluation, promoted officers drop out from the next evaluation. Since our data encompasses promotion outcomes recorded between 2014 and 2022, there will be some measurement error in recording the timing of the evaluation of officers. Nonetheless, it is important to control for the number of times we observe officers at the same rank re-evaluated for promotion in our period of analysis because the probability of promotion declines with each re-evaluation. As we describe below, we control flexibly for age, which captures many of the measurement error concerns.<sup>3</sup>

In the regression equation, we model promotion as a function of evaluation characteristics on the officer's most recent FITREP, before their promotion evaluation. Here, we include the difference between an individual's trait average on this most recent FITREP and their reporting senior's cumulative average (RSCA). We refer to this term as,  $\Delta$  (Trait Avg - RSCA), in the discussion below and in the tables. This delta term captures both the extensive margin of the evaluation, whether an individual's average is above or below their reporting senior's cumulative average (positive numbers lie above RSCA and negative numbers lie below RSCA), and the intensive margin (i.e., the degree to which an individual's trait average lies above or below their RSCA).

Since selection boards interpret an individual's trait average alongside their RSCA,  $\Delta$  (Trait Avg - RSCA) captures both aspects of the trait evaluation. While an individual's trait average can affect  $\Delta$  (Trait Avg - RSCA), their evaluation is also affected by whether their reporting senior has a high or low RSCA. Reporting seniors with lower RSCAs (i.e., harsh evaluators) can give exceptional officers very high trait averages compared to their RSCA (large positive deltas), while reporting seniors with higher RSCAs (i.e., easy evaluators) may be unable to signal exceptional candidates to the boards (small positive deltas) because the trait average and RSCA are capped at 5. For example, an officer with a 4.8 trait average will be evaluated more favorably if their reporting senior's cumulative average is 3.8 ( $\Delta$  (Trait Avg - RSCA) = 1) compared to an officer with the same 4.8 trait average but with a reporting senior whose RSCA is 4.5 ( $\Delta$

---

<sup>3</sup> As a robustness check, we estimated the model for individuals with their first evaluation in 2015 or higher because our assumption is that officers will be looked at every year (once they are in-zone) and there should not be gap years in board outcomes. This robustness check ensures we are analyzing such individuals. Our main results are essentially unchanged.

(Trait Avg - RSCA = 0.3)). As a robustness check, we introduce the individual trait average and RSCA as separate predictors to highlight the effect of  $\Delta$  (Trait Avg - RSCA) on promotion.

In addition to the trait average and RSCA, we include indicators for individuals that receive an EP and an MP, and for being in a non-competitive summary group (1 of 1). We control for the “1 of 1” group with an indicator variable instead of the summary group size because using summary group size (in another model specification) yielded a small positive effect, implying that being evaluated with more peers leads to a higher likelihood of promotion. However, we found that this effect is almost entirely driven by the negative impact of being in a non-competitive summary group. Once this factor is accounted for, the size of the group becomes immaterial to promotion outcomes.

The omitted groups are Promotable (and the very small number of those with Progressing and Significant Problems) and individuals in competitive summary groups. We also test for differential effects of receiving an EP and MP based on the summary group by including interactions of EP and MP with an indicator for 1 of 1. The total effect of receiving an EP in a non-competitive summary group is the sum of the regression coefficients on EP, 1 of 1, and the interaction of 1 of 1 with EP. The same interpretation holds for the total effect of receiving an MP in a non-competitive summary group. In the next section, we calculate these effects explicitly for each rank.

Finally, we also control for the number of evaluations completed by an officer’s reporting senior on their most recent FITREP. This captures the seniority and experience of reporting seniors that can influence promotion outcomes directly through well-executed FITREPS and indirectly by improving the performance of their subordinates that is then captured in their FITREP assessment.

Apart from evaluation characteristics, we also construct measures of the number of AQDs noted on the most recent evaluation and the number of awards earned over the last five years, along with an indicator for officers with a graduate education and those that are prior enlisted. We anticipate graduate education would increase the probability of promotion, while prior enlisted status may decrease the probability due to other uncaptured socio-demographic or professional characteristics. In terms of demographics, we include indicators for females, African Americans, Non-Black minorities, and married

officers. We also control for the number of dependents as of the year in which the candidate is considered for promotion.

In principle, gender and race indicators should not impact promotion in an unbiased evaluation system unless they are correlated with other factors that affect promotion, (for example, prestige of duty station, importance of the billet/assignment, etc.) and are omitted from our model (i.e., omitted variables bias). After controlling for all other factors correlated with gender and race, if race and gender are significant and consistent predictors of promotion, that may be evidence of conscious or unconscious bias in the selection board process.

It is worth clarifying that two potential sources of race or gender discrimination can be introduced at different times. First, reporting seniors may be biased, and they may give a subordinate of a particular race or gender lower-than-deserved trait scores. Second, promotion board members may be biased, and they may decline to promote a candidate with a worthy record. Our model takes the evaluations as given and does not directly explore if evaluations themselves are biased along gender and race lines. Rather, we model the decision-making process of board members when they observe an officer's evaluation and career characteristics.

Under these assumptions, if demographic characteristics are substantively and consistently correlated with promotion outcomes, there are three possible causes. First, the gender/race of the candidate truly impacts board decisions: board members and/or briefers choose to favor/disfavor candidates based on their demographic characteristics. This could arise from conscious or unconscious bias.

Second, gender/race of a candidate and promotion may be correlated with some important but omitted factors. For example, if white/male candidates are more likely to have graduated with a STEM degree, and if board members believe STEM skills are important for leadership positions, they may choose to promote more candidates with STEM degrees. Another possible channel may be through shared, common experiences between the candidate and board participants. If white/male candidates have unit identification codes (UIC) and billet assignments that are similar to board members' careers, they may feel more confident about promoting these candidates. Since our model does not include information on the college major of the candidate or the background of

board members because these data were unavailable, the model estimates may make it appear *as if* they favor white/male candidates. It should be emphasized that this is *not* discrimination.

Finally, if a gender/race is under-represented in the sample, we may isolate a *statistically significant* correlation that may be driven by “small sample” issues. For example, assume there are 100 candidates being considered at a board, and African American candidates comprise 10% of the sample. Further assume that 80% of candidates have records that merit promotion. Then, if race is completely uncorrelated with promotion, we should expect to see 8 (out of 10) African American candidates promoted. However, if one of these candidates is not promoted, we would measure a 12.5% disadvantage for African American compared to their white colleagues. This may be due to discrimination, a mistake, or a fluke. Because there are so few African American candidates, one mistake or fluke can yield a large estimated impact.

In addition to demographics and the most recent FITREP, we also capture an individual’s past performance records because performance summary reports for officers include information on past FITREPs that a briefer may choose to highlight at a selection board. Our past evaluation characteristics include an average of EPs received, MPs received,  $\Delta$  (Trait Avg – RSCA), number of evaluations completed by past reporting seniors, average of awards and AQDs received across the last four evaluations before the most recent evaluation. In cases where the past evaluation was missing, we include a zero for the past evaluation characteristics and a separate indicator to capture the missing evaluations.

While the factors described above capture the main predictors of promotion in our model, we also include a series of indicator variables to flexibly control for other individual factors that may affect officer promotion. First, we include indicators for an officer’s age, which allows age to affect promotion in a linear or non-linear fashion, whatever the case may be. Second, we include indicators for the two-digit MOS code of an officer to ensure we are comparing officers within the same MOS, and not across MOS. Third, we include indicators for the selection board year, which captures any annual idiosyncrasies that may affect promotion. Fourth, we include indicators for whether it is an officer’s first, second, third or fourth evaluation at a board. These

indicators, also known statistically as fixed effects for the different categories, ensure we are comparing apples-to-apples to the extent possible. To further explore differences by MOS, we estimate split-samples of the model by URL and RL-Staff, and then the large communities within URL (Aviation and SWO), and RL-Staff (aggregate group of MED-JAG).

## IV. RESULTS

We review the results separately by rank as our regressions are estimated separately by officer rank from O-3 candidates to O-5 candidates.

### A. O-3 TO O-4 PROMOTIONS

Table 4 presents the results for the 18,732 O-3 candidates evaluated by promotion boards between 2014 and 2022. Each column represents an individual regression with column (1) using the full sample of O-3 candidates, column (2) focusing on the URL community, column (3) on Aviation officers, column (4) on SWOs, column (5) on the RL-Staff community, and column (6) on MED-JAG officers among the RL-Staff. We report mean promotion in each analysis sample on the first row for ease of interpretation. Looking down the column, we organize the predictors into characteristics on the most recent evaluation, career characteristics, demographics, and past evaluation and career characteristics.

As seen in column (1), the coefficient on  $\Delta$  (Trait Avg – RSCA) on the most recent evaluation is large and positive indicating that individuals whose trait averages are larger than their reporting senior’s cumulative average are more likely to promote. In terms of standardized effect sizes, the coefficient translates into an effect size of 0.12 standard deviations (SD), which can be interpreted as a one standard deviation increase in  $\Delta$  (Trait Avg - RSCA) increases the probability of promotion by 0.12 SD. Another way to interpret the coefficient is that a one-point increase in  $\Delta$  (Trait Avg - RSCA) increases the probability of promotion by 15 percentage points against a mean of 69%. While the effect size varies across the different samples in columns (2) through (6), it is always large, positive, and statistically significant at the 1% level.

Receiving an EP on the most recent evaluation is another key predictor of promotion with a large and positive coefficient significant at the 1% level across the different samples. Receiving an EP increases the probability of promotion to O-4 by 22 percentage points (column 1), which is a large economic effect representing 32% of the mean of promotion. The coefficient on EP is remarkably stable across the split samples ranging from an effect size of 26% ( $=0.19/0.72$ ) of the promotion mean for MED-JAG officers in column (6) to 32% ( $=0.27/0.84$ ) of the promotion mean for SWOs in column

(4). Moreover, EP has a larger standardized effect size than  $\Delta$  (Trait Avg - RSCA) at 0.23 SD compared to 0.12 SD (using column [1] estimates).

Unlike  $\Delta$  (Trait Avg - RSCA) and EP, receiving an MP increases the probability of promotion by 7 percentage points (standardized effect size of 0.06 SD), while being a 1 of 1 (i.e., a non-competitive summary group) reduces the probability of promotion by 6 percentage points (standardized effect size of 0.04 SD) using column (1) coefficients. We find no differential additional effect of receiving an EP or an MP in non-competitive summary groups because the coefficients on the interaction of 1 of 1 with EP and MP are both statistically insignificant. Note, the *total effect* of receiving an EP in a non-competitive group is smaller than receiving an EP in a competitive group. Using estimates from column (1), compared to the reference group of individuals that received a Promotable (or below) in a competitive group (not a 1 of 1), the *total effect* of receiving an EP as a 1 of 1 increases the probability of promotion by a statistically significant 12 percentage points ( $=0.2213-0.0586-0.0407$ ). In contrast, the total effect of receiving an EP in a competitive group increases the probability of promotion by 22 percentage points as noted above. For MPs, the *total effect* of receiving an MP in a non-competitive summary group reduces the probability of promotion by 3 percentage points ( $=0.0735-0.0586-0.0497$ ) compared to the reference group, but the effect is not statistically significant indicating that receiving an MP in a non-competitive summary group does not statistically affect the probability of promotion compared to individuals receiving a P in competitive groups. Finally, the coefficient on the number of evaluations completed by reporting seniors is statistically significant, but it is not economically significant because of the small effect sizes.

Among career characteristics, both the number of awards and AQDs received by an officer increase the probability of promotion as does graduate education. Of the three, AQDs have the largest standardized effect size at 0.15 SD compared to 0.05 SD for awards and 0.04 SD for graduate education (using column (1) coefficients). In terms of demographics, being married has the largest standardized coefficient at 0.04 SD, which is significantly smaller than the coefficients on evaluation and career background variables. Being female increases the probability of promotion to O-3 by 1.3 percentage points albeit not at conventional levels of significance, while being African American or a Non-

Black minority reduces the probability of promotion by 2.2 and 2.9 percentage points, respectively. These translate into standardized coefficients of 0.01 SD and 0.02 SD, relatively small effect sizes. We explore potential explanations for the race and gender results further in Section D on high and low performers.

Past evaluation characteristics are also predictors of promotion with the biggest predictor being the average of the historical delta between an individual's trait average and their RSCA across their past evaluations (standardized effect size of 0.15 SD). Past AQDs and awards are less significant. Comparing the different officer characteristics, a large and positive  $\Delta$  (Trait Avg - RSCA) along with receiving an EP on the most recent FITREP are the two most important predictors of promotion to O-4.

Looking across the columns,  $\Delta$  (Trait Avg - RSCA) and an EP are uniformly predictive of promotion across the URL and RL-Staff and within the larger URL and RL-Staff communities. The coefficients on MP are consistently smaller, and in the case of Aviation, in column (3) insignificant suggesting MPs are evaluated less favorably in Aviation compared to the other communities. Being a 1 of 1 hurts URL officers and especially those in Aviation (column [3]) but has no significant effect among SWOs (column [4]). However, receiving an EP as a 1 of 1 candidate reduces the probability of promotion by 11 percentage points in the RL-Staff community effectively cutting the EP effect almost in half compared to an RL-Staff officer receiving an EP in a competitive category (not a 1 of 1). Compared to the reference group of RL-Staff individuals with a P in competitive groups, the *total effect* of receiving an EP in a competitive group increases the probability of promotion by 22 percentage points, while the *total effect* of receiving an EP in a non-competitive group increases the probability of promotion by only 13 percentage points (column [5] coefficients).<sup>4</sup>

Receiving awards and AQDs consistently increases the probability of promotion, albeit less for the MED-JAG. Graduate education is a key predictor for the RL-Staff, but less so for the URL, especially those in Aviation. Being female increases the probability of promotion among RL-Staff, as does being married. In contrast, being female reduces

---

<sup>4</sup> RL-Staff Corps encompasses many different communities, and the effects of competitive/non-competitive summary groups may be heterogeneous across individual communities. Unfortunately, the sample sizes are too small to investigate heterogeneity within individual RL-Staff communities apart from the larger MED-JAG group that we analyze in column (6).

the probability of promotion in Aviation. The negative effects of race in column (1) for the full sample are largely driven by the negative effects of being African American and Non-Black in Aviation. We do not find any significant negative coefficients on the race variables among SWOs or the RL-Staff.

Given the significant effects on  $\Delta$  (Trait Avg - RSCA) in the most recent evaluation, we present results in Appendix Table 1 where we include an individual's trait average and RSCA as separate predictors in lieu of the difference between the two to highlight their individual roles in affecting officer promotion. Having a higher trait average controlling for the RSCA increases the probability of promotion as we would expect. But, as discussed earlier, an officer with a more lenient evaluator as a reporting senior (i.e., someone with high RSCAs) reduces an officer's probability of promotion. This accounts for the negative coefficient on the reporting senior's cumulative average.

In our view, this is a serious problem with the current system that is only partially corrected by promotion recommendations because the trait averages and RSCA continue to affect promotion even after controlling for the promotion recommendation. Officers under reporting seniors with high RSCAs are unable to stand out even with high individual trait averages and EPs because their reporting senior is unable to signal their exceptional quality. While this hurts top performers, it does not substantively impact lower performers. Although they may receive generous scores under their lenient boss, the high RSCA value will lead the promotion board to discount these scores appropriately. We should note that with Navy FITREPs suffering from grade inflation (average trait scores across all ranks is approximately 4.2), moving to a disciplined reporting senior with a RSCA value (of, say, 3.5) would increase the likelihood of promotion of an exceptional candidate (who receives an average score of 5) by about 17 percentage points as compared to a RS with a RSCA of 4.5.

## **B. O-4 TO O-5 PROMOTIONS**

Table 5 summarizes the regressions results for O-4 candidates up for promotion to O-5. The columns corresponding to the split samples are organized in the same manner as in Table 4. Like O-3 candidates, the coefficient on  $\Delta$  (Trait Avg - RSCA) on the most recent evaluation is again large and statistically significant representing a standardized effect size of 0.14 SD. Receiving an EP increases the probability of promotion to O-5 by

25 percentage points against a mean of 58%, i.e., an effect size of 44% (standardized effect of 0.23 SD, column [1] estimates). This is an incredibly large effect size ranging from 24% of mean promotion among Aviation officers (column [3]) to 53% for MED-JAG (column [6]). O-4 candidates in non-competitive categories (i.e., 1 of 1) are penalized only when they receive an EP (column [1]). In this case, receiving an EP and being a 1 of 1 increases the probability of promotion by only 14 percentage points ( $=0.2527+0.0602-0.1717$ ) compared to 25 percentage points for O-4 candidates in competitive categories relative to the reference group of individuals with P in competitive categories. The differential effects of receiving an EP and being a 1 of 1 are especially negative for the URL and MED-JAG community. Given the differences across MOS in summary group sizes, it is unsurprising there is heterogeneity by MOS in the effects of competitive categories on promotion.

Looking down the regression results for the full sample in column (1), awards are less economically significant than for O-3 candidates, each additional award or AQD increases the probability of promotion by 2.2 percentage points while having a graduate education increases the probability of promotion by 6 percentage points. Being female is uncorrelated with promotion to O-5, but being African American or a Non-Black minority reduces the probability of promotion by 4 percentage points on average, an economic effect of -0.02 SD for African Americans and -0.03 SD for Non-Black minorities. Unlike O-3 candidates where the negative effects on race were driven by Aviation, in the case of O-4 candidates the negative effects for African Americans are driven by SWOs and for Non-Black minorities by MED-JAG. Among past evaluation characteristics, the biggest predictor of promotion is an officer's historical average of  $\Delta$  (Trait Avg - RSCA) with an economic effect of 0.18 SD. It is also predictive of promotion across communities as seen in columns (2) through (6).

Looking across communities, we observe some differences between the URL and RL-Staff where receiving an MP reduces the probability of promotion among the URL, especially those in Aviation, while an MP increases the probability of promotion in the RL-Staff albeit by significantly less than receiving an EP. For example, among MED-JAG officers, receiving an EP increases the probability of promotion by 29 percentage points compared to 6 percentage points for an MP relative to those receiving P (the

omitted group). Awards and AQDs are important predictors for O-4 candidates in Aviation but are less consistently predictive among SWOs (awards) and MED-JAG (AQDs).

Appendix Table 2 shows the results when we split the  $\Delta$  (Trait Avg – RSCA) into two variables, namely an individual’s trait average and their reporting senior’s cumulative average. We again find large positive effects on the individual trait average along with negative effects on RSCA highlighting the problem of strict versus more generous evaluators among reporting seniors. Broadly, the results on O-4 match the patterns we observed among O-3 candidates in that  $\Delta$  (Trait Avg – RSCA) and EP significantly increase the probability of promotion, but there are differences by rank on the effects of competitive categories and career characteristics.

### **C. O-5 TO O-6 PROMOTIONS**

In Table 6, we present results for O-5 candidates being considered for promotion to O-6. The table is organized in the same manner as Tables 4 and 5. Because of the smaller sample, we are cautious drawing strong conclusions for this group. This is especially true for communities like SWOs where we only observe 781 officer\*board year observations between 2014 and 2022. In the case of O-5 candidates, we control for characteristics of the most recent FITREP, career characteristics, and demographics.

In contrast to our analyses of lower ranks, we do not control for past evaluation and career characteristics for O-5 candidates. Individuals that have attained a rank of O-5 are already a highly selected group with above average past evaluations. Only about 40% of O-3 officers will be promoted to the rank of O-5. Since time-in-grade requirements for O-4 and O-5 are 3 years each, using the last 5 evaluations may include information that was used to trim the sample twice in a deliberate manner (to isolate only high performers). Since the analysis is already conditioning on the past evaluations of O-5 candidates, the analysis would generate severely biased estimates of these variables.

Similar to the other ranks, the coefficient on  $\Delta$  (Trait Avg – RSCA) is large, positive and statistically significant at the 1% level across communities. In the case of the full sample (column [1]), the coefficient of 0.2541 corresponds to an economic effect of 0.19 SD. While EP continues to play a role in promotion, it is no longer a statistically significant predictor of promotion among SWOs (column [4]). Interestingly, the

experience of reporting seniors as captured by the number of completed evaluations is positive and significant though the effect sizes are relatively small. Nonetheless, it perhaps highlights the importance of experienced mentors at more senior ranks.

Awards and AQDs continue to play a role. Graduate education increases the probability of promotion to O-6 in column (1), but there are marked differences by community with graduate education being a more important predictor for RL-Staff and MED-JAG. Unlike the lower ranks, gender and race play a different role at the senior rank. Being female is uncorrelated with promotion in most cases, apart from column (1) where it is has a positive but only marginally significant effect. Being African American is also uncorrelated with promotion for the URL and positively correlated with promotion among the RL-Staff and MED-JAG. In contrast, being a Non-Black minority reduces the probability of promotion among the URL, especially among SWOs.

#### **D. DISCUSSION OF OFFICER PROMOTION RESULTS**

Broadly, the results on officer promotion highlight the importance of the most recent FITREP on promotion outcomes. Across rank and community, individuals with bigger deltas between their individual trait average and their reporting senior's cumulative average and an EP are more likely to promote holding all else equal. Unlike these two variables, the effect of less competitive summary groups (1 of 1) is more nuanced with differential effects by rank and community. This points to variation at selection boards in the interpretation of summary groups. Aviation officers in non-competitive summary groups are at a disadvantage compared to SWO or those in the RL-Staff community.

Past awards play a big role among O-3 and O-4 Aviation candidates while graduate education plays a smaller role in affecting promotion outcomes in the Aviation community. In contrast, graduate education plays a bigger role among the RL-Staff across ranks. This is as we would expect, given in many RL-Staff communities, such as MED-JAG, a graduate or professional education is necessary to perform the job duties (e.g., a Navy doctor or lawyer).

Unlike the results on evaluations and officer career characteristics, the results on demographics are more difficult to interpret. In theory, race and gender can affect promotion for three mutually nonexclusive reasons as discussed earlier. First, race and

gender may be correlated with other unobservable factors that affect promotion such as the importance of the past duty location of an officer or Aviation Selection Test Battery scores in the case of aviators. In this case, the race and gender variables may pick up the effects of these omitted factors. Second, race and gender results may be indicative of discrimination (positive or negative) at the selection boards due to conscious or unconscious bias among selection board members. In this case, selection boards could be using stereotypes to infer future potential, which in turn hurts certain demographics. In economics, a lack of information drives such statistical discrimination. Such discrimination is different from taste or preference-based discrimination where animosity towards certain demographics can lead to worse outcomes for that group be it promotion to the next rank or wages.<sup>5</sup> Three, if the race and gender subgroup comprise a small fraction of the entire sample, the econometric model can return a statistically significant result arising out of a small number of personnel mistakes or fluke incidents.

If the negative results we observe on race for O-3 and O-4 candidates are driven by systematic bias against these groups, we would expect to observe uniformly negative coefficient on indicators for African American and non-Black minority officers. Yet, we do not observe that to be the case. African American O-3 candidates in the Aviation community are slightly less likely to promote as are Non-Black minorities compared to white candidates. We find no such negative effects among SWOs or the RL in the O-3 candidate pool. Among O-4 candidates, race does not affect promotion in the Aviation community but African American O-4 candidates in the surface warfare community are less likely to promote. Finally, we observe no negative effects for African American O-5 candidates. If anything, African American O-5 candidates are more likely to promote in the RL-Staff community.

One way to test for systemic taste-based discrimination is to consider the records of exceptionally high or low performers to assess if demographics affect promotion in these groups. If we find that race or gender consistently play an outsized role among candidates that are clearly superior or inferior to the average candidate, this would serve as a “smoking gun” and indicate significant bias in the promotion process. This bias

---

<sup>5</sup> See Anderson, Fryer, and Holt (2006) on an economics overview on identifying discrimination.

would be glaringly obvious, large enough to show unfit candidates being promoted over excellent candidates due to race or gender.

To test this conjecture, we define high performers as individuals who have only received EPs across their observed FITREPS including the most recent and low performers are individuals that have never received an EP across their observed FITREPs. This clearly and severely splits the sample. For example, among O-4 officers, the high performers have a 68% chance of promotion, while the low performers have a 16% chance of promotion.

Table 7 presents the results disaggregated by rank. None of the race variables are statistically significant in this group of high and low performers. Moreover, the results on gender are indicative of positive discrimination at the higher ranks. High performing females are 7.5 percentage points more likely to promote to O-6 compared to their male counterparts. The female effect is also visible among low performers in the O-5 candidate pool. An individual's assessment by their reporting senior appears to be the key predictor of promotion of high performers across rank, while being in a non-competitive summary group uniformly decreases the probability of promotion among high performers. While these results do not rule out *all* bias in the evaluation process, they suggest there is no evidence of obvious bias among the top and bottom performers. Systemic taste-based discrimination is thus an unlikely driver of the results on race in Tables 4 and 5.

#### **E. ENLISTED RANKS**

As described in the data section, unlike officers where we have better data, we are unable to study “promote” or “fail to promote” as an indicator outcome for enlisted personnel. The promotion status for enlisted sailors in the inventory data shared by NAVPERS never changes from “not considered for promotion,” and thus, we can only use the change in enlisted ranks to calculate time to promotion.

Time to promote is a very noisy measure of promotion because it excludes individuals that are never promoted and does not perfectly capture the time to promotion due to the issues discussed above. Nonetheless, we show the results for completeness but hesitate to draw strong conclusions because the outcome (time to promote) is only measured for individuals that promote.

In the case of enlisted sailors, we restrict the analysis to time to promotion for enlisted sailors of rank E-5 that are promoted to rank E-6. This is the largest enlisted group in our personnel data sample. Time to promote is also measured with less error for this rank than higher enlisted ranks because of increasing requirements. Because of limitations in the data, we cannot tell if an E-6 sailor who did not promote to E-7 upon meeting the minimum time in rating requirement failed to be selected or did not yet complete all requirements.<sup>6</sup>

Using a similar regression model as officers, in Table 8 we model time to promotion as a function of a Sailor's most recent evaluation characteristics, their career characteristics, and demographics. In these regressions, a negative coefficient indicates those variables shorten the time to promotion, while positive coefficients indicate a longer time to promotion. On average, E-5 Sailors promote to E-6 in 53 months, or just under 5 years. Considering the full sample of enlisted E-5 sailors that promoted to E-6 in column (1), the results on the evaluation characteristics are as we would expect. Individuals with large deltas between their individual trait average and their reporting senior cumulative average promote faster as do individuals that received an EP while, those in non-competitive summary groups promote three months later all else equal. In contrast to the relatively small effect sizes for evaluation characteristics, education is a large predictor of time to promotion for enlisted personnel. Individuals with a college education promote 8 months sooner than those without a college education and those with a graduate education promote 23 months sooner than those without any college education. With regard to demographics, African Americans and non-Black minority sailors promote faster than white sailors though the effect sizes are small.

In column (2), we disaggregate the delta variable into the individual trait average and the reporting senior cumulative average. Similar to the officer results, individuals with a higher individual trait average promote 3 months sooner but Sailors under more lenient reporting seniors (i.e., those with higher RSCAs) also promote faster. Unlike officers, enlisted sailors with exceptional records are not penalized if their reporting

---

<sup>6</sup> For promotion to E7, beyond basic requirements common across promotion to E4 to E6, the sailor must complete technical training through class A schools, the Chief Petty Officer (CPO) Indoctrination course, pass the Navy Advancement Exam, and be evaluated by a CPO selection board.

senior has a higher RSCA. This may be related to lower trait averages among enlisted personnel (4.05 for E-5 promoted to E-6 compared to 4.24 for O-3 candidates). In column (3), we look at only Administration, Deck, Technical, and Weapons Specialty ratings. This group comprises the biggest set of sailors and excludes smaller, more specialized ratings, such as Medical or Engineering ratings. In column (4) we restrict the analysis to individuals that were promoted in 2018 or earlier. As our evaluation data begins in 2014, promotions past 2018 means we are retaining sailors who may take 6 years or more to promote to E-6. Those taking at least twice as long to promote as their peers may be fundamentally different and may contaminate the econometric results.

We find similar results as for the main sample in column (1). In conclusion, the enlisted analysis suggests education and career background characteristics are larger drivers of time to promotion compared to evaluation characteristics, and women and minorities promote faster than men and white Sailors.

## V. CONCLUSION AND RECOMMENDATIONS

Using detailed data on officer evaluation characteristics, career characteristics and demographics, we find average trait scores, RSCA, and the EP promotion recommendation on the most recent FITREP/EVAL are highly predictive of promotion between 2014 and 2022 across O-3, O-4, and O-5 promotion candidates. Earning an average trait score above RSCA and an EP are most likely to lead to promotion across ranks and communities. Yet, there are substantive differences in promotion and drivers of promotions by community and rank highlighting heterogeneity in the selection process for officers by community. For example, the impact of an EP (compared to the baseline of Promote) ranges from 18 percentage points (Aviation O-3) to 27 percentage points (SWO O-3) among O-3 candidates. We observe similar differences by MOS for O-4 and O-5 candidates. Career characteristics such as awards, AQDs and graduate education also have differential effects by rank and community.

With regard to demographics, women and minorities are under-represented with larger under-representation at higher ranks. For example, women make up 18% of promotion candidates from O-3 to O-4, but they only account for 13% of O-5 to O-6 candidates. Similarly, while non-white officers make up 27% of O-3 to O-4 promotion candidates while they account for 22% of O-5 to O-6 candidates. In the regression analysis, we find mixed results for socio-demographic characteristics. Being female or married is sometime positively associated with promotion, while being African American or a non-Black minority is sometimes negatively associated with promotion. These results on race and gender are economically smaller than the evaluation characteristics and are also inconsistent across ranks and communities. We fail to find “smoking guns” of bias by looking at promotions of high and low performers. High performers have all EPs and low performers never have an EP across all observable evaluations. In these groups, race and gender are mostly insignificant drivers of promotions suggesting widespread and systemic bias are perhaps not driving the race results.

Looking from O-3 to O-5 candidates, promotions become less predictable at higher ranks. While the impacts described qualitatively hold for higher ranks, their magnitude and statistical significance fluctuates, such as the effect of receiving an EP.

This finding, along with the mixed results for socio-demographic characteristics strongly imply that we are missing important factors that drive promotions. We suspect that information on the officer summary report, narrative summaries on the FITREPs, ability and experience of selection board briefers, and composition of the selection board may affect promotion and are uncaptured in the statistical analysis.

Based on our findings, we offer a series of recommendations with regard to process and future research. First, we recommend the Navy mandate a ceiling for RSCA so that high performers working under reporting seniors with high RSCAs are not disadvantaged. Average trait score above RSCA is one of the most salient predictors of promotion; however, a high RSCA restricts the ability of high performers to “stand out.” While our analysis does not generate a specific recommended value for the ceiling, at the very least we recommend an average ceiling for RSCA below 4.0 to ensure reporting seniors can credibly signal their high performers.

Second, we recommend the Navy publicize promotion rates and factors positively associated with promotion to all participants in the promotion process, including the sailors and officers being evaluated, reporting seniors, briefers, and board members. In particular, differences across ranks and communities should be clarified, and factors that play a large (and not-so-large) role in promotion should be explained so that all participants have the most complete and accurate information.

Third, we recommend future research to clarify the mixed impacts on socio-demographic characteristics found in this study. Cognitive biases may exist during Navy selection boards. To ensure the promotion of the best and brightest candidates, the Navy needs to evaluate whether biases affect the evaluation process and identify strategies to mitigate them. In particular, we recommend the following research questions for future research to explore:

- Do the background characteristics of the Board President and members (gender, race, academic, professional background, community) affect the promotion of officers with characteristics similar to or different from Board members?
- Do the timing of the evaluation of records (reviewed later in the day or selection window) and fatigue of Board participants affect promotion?

- Do the background characteristics and briefing experience/performance of selection board briefers affect promotion?
- How can we mitigate such biases if they do exist?

Finally, the data analysis in this project was hampered with an inability to track key evaluation metrics that are not systematically recorded in the NAVPERS data. To alleviate these issues in future research, we recommend that NAVPERS augment their data collection with more precise information, especially for the inventory dataset. Specifically, zone designation, written comments, timing of the briefs at the promotion board (to account for fatigue of briefers and board members), the identity of the briever as well as board members should be collected and maintained. We also encourage the Navy to allow researchers to visit and observe selection boards to collect data on the premises, especially on how briefers convey positive or negative information about the candidate and how fatigue of board members/briefers and experience/expertise of the briever may impact outcomes. Such an evaluation will complement the statistical evaluation of retrospective data. Both efforts will enable the Navy to achieve a more fair, effective, and efficient Performance Evaluation system.

## TABLES

**Table 1A: Candidates for Promotion O-3 to O-4**

Boards between 2014 and 2022; 18,732 individuals with 1.3 boards per individual

	Full Sample	MOS URL	MOS Aviation	MOS Surface Warfare	MOS RL + Staff Corp	MOS Med-Jag
Promote	69%	72%	64%	84%	67%	72%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.20	0.20	0.25	0.12	0.21	0.19
Trait Avg.	4.24	4.27	4.32	4.25	4.21	4.06
EP	65%	63%	59%	65%	67%	59%
MP	24%	26%	27%	27%	22%	28%
1 of 1	33%	26%	28%	23%	39%	19%
<i>Officer Career Characteristics</i>						
Awards (#)	4.83	5.34	6.02	3.75	4.41	4.27
AQDs (#)	3.49	3.59	3.05	4.83	3.40	1.59
MOS URL	46%	100%	100%	100%	0%	0%
MOS Aviation	27%	60%	100%	0%	0%	0%
MOS Surface Warfare	12%	25%	0%	100%	0%	0%
MOS Medical/JAG	23%	0%	0%	0%	42%	100%
Graduate Educ (0/1)	40%	34%	22%	59%	45%	61%
Prior Enlisted	12%	6%	5%	8%	16%	8%
<i>Demographics</i>						
Female	18%	10%	9%	16%	25%	39%
Black	9%	5%	3%	11%	12%	10%
Non-Black Minority	18%	15%	13%	19%	21%	22%
Married	77%	77%	76%	73%	78%	76%
Dependents (#)	1.29	1.08	1.00	1.16	1.47	1.27
Age	35.49	33.39	33.06	33.99	37.25	36.34
Observations	21,711	9,889	5,887	2,520	11,822	4,992

**Table 1B: Candidates for Promotion O-3 to O-4**

Boards between 2014 and 2022; 18,732 individuals with 1.3 boards per individual

	Full Sample	Male	Female	White	Black	Non-Black Minority
Promote	69%	69%	71%	71%	61%	65%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.20	0.20	0.22	0.21	0.17	0.18
Trait Avg.	4.24	4.25	4.20	4.25	4.22	4.21
EP	65%	66%	64%	65%	66%	64%
MP	24%	23%	25%	24%	22%	24%
1 of 1	33%	34%	30%	32%	42%	35%
<i>Officer Career Characteristics</i>						
Awards(#)	4.83	4.95	4.29	4.95	4.17	4.67
AQDs (#)	3.49	3.65	2.74	3.46	3.65	3.51
MOS URL	46%	50%	24%	50%	25%	37%
MOS Aviation	27%	30%	13%	31%	9%	20%
MOS Surface Warfare	12%	12%	10%	11%	15%	12%
MOS Medical/JAG	23%	17%	50%	22%	26%	27%
Graduate Educ (0/1)	40%	38%	48%	39%	43%	41%
Prior Enlisted	12%	12%	10%	10%	22%	15%
<i>Demographics</i>						
Female	18%	0%	100%	16%	28%	20%
Black	9%	8%	14%	0%	100%	0%
Non-Black Minority	18%	18%	20%	0%	0%	100%
Married	77%	81%	60%	79%	69%	76%
Dependents (#)	1.29	1.40	0.79	1.24	1.62	1.31
Age	35.49	35.64	34.83	35.07	37.92	35.99
Observations	21,711	17,811	3,900	15,782	1,966	3,963

**Table 2A: Candidates for Promotion O-4 to O-5**

Boards between 2014 and 2022; 9,325 individuals, 1.5 boards per individual

	Full Sample	MOS URL	MOS Aviation	MOS Surface Warfare	MOS RL + Staff Corp	MOS Med-Jag
Promote	58%	62%	60%	59%	55%	54%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.20	0.24	0.34	0.11	0.18	0.16
Trait Avg.	4.41	4.49	4.47	4.57	4.34	4.24
EP	72%	72%	70%	66%	72%	66%
MP	16%	16%	18%	18%	17%	20%
1 of 1	37%	35%	33%	38%	39%	21%
<i>Officer Career Characteristics</i>						
Awards (#)	4.44	5.42	6.27	4.27	3.69	3.51
AQDs (#)	4.65	4.86	4.87	5.30	4.48	2.12
MOS URL	44%	100%	100%	100%	0%	0%
MOS Aviation	23%	54%	100%	0%	0%	0%
MOS Surface Warfare	13%	29%	0%	100%	0%	0%
MOS Medical/JAG	24%	0%	0%	0%	42%	100%
Graduate Educ (0/1)	69%	60%	51%	77%	76%	75%
Prior Enlisted	1%	1%	0%	0%	1%	1%
<i>Demographics</i>						
Female	15%	6%	4%	13%	22%	34%
Black	9%	6%	3%	14%	11%	11%
Non-Black Minority	15%	13%	11%	16%	17%	18%
Married	86%	88%	89%	85%	85%	84%
Dependents (#)	1.84	1.77	1.74	1.70	1.88	1.84
Age	40.95	39.13	38.71	39.80	42.36	42.13
Observations	11,593	5,046	2,707	1,465	6,547	2,767

**Table 2B: Candidates for Promotion O-4 to O-5**

Boards between 2014 and 2022; 9,325 individuals, 1.5 boards per individual

	Full Sample	Male	Female	White	Black	Non- Black Minority
Promote	58%	58%	57%	60%	49%	53%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.20	0.21	0.18	0.22	0.14	0.17
Trait Avg.	4.41	4.42	4.35	4.42	4.38	4.37
EP	72%	72%	69%	73%	67%	69%
MP	16%	16%	19%	16%	19%	17%
1 of 1	37%	38%	32%	37%	40%	37%
<i>Officer Career Characteristics</i>						
Awards (#)	4.44	4.60	3.54	4.53	3.95	4.30
AQDs (#)	4.65	4.82	3.64	4.68	4.32	4.67
MOS URL	44%	48%	17%	47%	30%	37%
MOS Aviation	23%	26%	6%	26%	8%	17%
MOS Surface Warfare	13%	13%	11%	12%	19%	13%
MOS Medical/JAG	24%	19%	54%	22%	30%	28%
Graduate Educ (0/1)	69%	68%	77%	68%	71%	74%
Prior Enlisted	1%	1%	1%	1%	1%	1%
<i>Demographics</i>						
Female	15%	0%	100%	14%	23%	17%
Black	9%	8%	14%	0%	100%	0%
Non-Black Minority	15%	15%	18%	0%	0%	100%
Married	86%	89%	69%	88%	79%	83%
Dependents (#)	1.84	1.94	1.26	1.85	1.94	1.73
Age	40.95	40.98	40.81	40.64	42.64	41.47
Observations	11,593	9,863	1,730	8,746	1,053	1,794

**Table 3A: Candidates for Promotion O-5 to O-6**

Boards between 2014 and 2022; 4,971 individuals, 1.5 boards per individual

	Full Sample	MOS URL	MOS Aviation	MOS Surface Warfare	MOS RL + Staff Corp	MOS Med-Jag
Promote	46%	46%	36%	53%	46%	47%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.08	0.04	0.02	0.06	0.12	0.13
Trait Avg.	4.49	4.55	4.55	4.65	4.44	4.37
EP	72%	69%	68%	73%	74%	65%
MP	13%	14%	15%	14%	12%	15%
1 of 1	45%	45%	44%	49%	44%	25%
<i>Officer Career Characteristics</i>						
Awards (#)	3.64	4.02	4.16	3.93	3.37	3.17
AQDs (#)	4.93	5.10	5.36	4.87	4.80	2.20
MOS URL	0.42	100%	100%	100%	0%	0%
MOS Aviation	0.22	53%	100%	0%	0%	0%
MOS Surface Warfare	0.12	29%	0%	100%	0%	0%
MOS Medical/JAG	0.26	0%	0%	0%	45%	100%
Graduate Educ (0/1)	84%	80%	73%	91%	87%	84%
Prior Enlisted	0%	0%	0%	0%	0%	0%
<i>Demographics</i>						
Female	13%	5%	3%	10%	20%	28%
Black	8%	5%	3%	9%	10%	8%
Non-Black Minority	14%	11%	9%	17%	16%	17%
Married	90%	93%	92%	92%	88%	86%
Dependents (#)	2.06	2.09	2.09	2.01	2.04	2.05
Age	46.03	44.59	44.31	45.07	47.07	47.13
Observations	6,368	2,653	1,408	781	3,715	1,687

**Table 3B: Candidates for Promotion O-5 to O-6**

Boards between 2014 and 2022; 4,971 individuals, 1.5 boards per individual

	Full Sample	Male	Female	White	Black	Non- Black Minority
Promote	46%	46%	45%	46%	48%	43%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.08	0.08	0.11	0.08	0.09	0.08
Trait Avg.	4.49	4.49	4.43	4.49	4.48	4.46
EP	72%	72%	68%	72%	74%	72%
MP	13%	13%	15%	13%	13%	13%
1 of 1	45%	46%	38%	44%	49%	46%
<i>Officer Career Characteristics</i>						
Awards (#)	3.64	3.70	3.24	3.66	3.46	3.61
AQDs (#)	4.93	5.05	4.12	4.81	5.77	5.10
MOS URL	42%	46%	15%	45%	25%	33%
MOS Aviation	22%	25%	5%	25%	8%	14%
MOS Surface Warfare	12%	13%	9%	12%	15%	15%
MOS Medical/JAG	26%	22%	56%	26%	26%	31%
Graduate Educ (0/1)	84%	84%	88%	83%	85%	88%
Prior Enlisted	0%	0%	0%	0%	0%	0%
<i>Demographics</i>						
Female	13%	0%	100%	12%	23%	17%
Black	8%	7%	13%	0%	100%	0%
Non-Black Minority	14%	13%	18%	0%	0%	100%
Married	90%	93%	71%	91%	81%	89%
Dependents (#)	2.06	2.16	1.42	2.10	1.91	1.93
Age	46.03	46.02	46.10	45.85	47.18	46.41
Observations	6,368	5,515	853	4,978	491	899

**Table 4: O-3 to O-4 Promotion, Outcome = Promote (0/1)**

	(1) Full Sample	(2) URL	(3) Aviation	(4) Surface Warfare	(5) RL + Staff Corp	(6) Med-Jag
Promote	69%	72%	64%	84%	67%	72%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.1454*** (0.009)	0.1128*** (0.012)	0.1179*** (0.016)	0.0941*** (0.027)	0.1758*** (0.015)	0.1358*** (0.025)
EP	0.2213*** (0.012)	0.2119*** (0.019)	0.1758*** (0.026)	0.2661*** (0.039)	0.2139*** (0.016)	0.1867*** (0.021)
MP	0.0735*** (0.012)	0.0883*** (0.018)	0.0126 (0.022)	0.2233*** (0.039)	0.0542*** (0.015)	0.0456** (0.019)
1 of 1	-0.0586** (0.028)	-0.1185*** (0.041)	-0.1557*** (0.044)	0.1127 (0.201)	0.0259 (0.039)	-0.0701 (0.138)
1of1*EP	-0.0407 (0.029)	-0.0004 (0.042)	0.0037 (0.046)	-0.1694 (0.201)	-0.1074*** (0.039)	0.0196 (0.139)
1of1*MP	-0.0497 (0.036)	-0.0583 (0.051)	0.0197 (0.055)	-0.1216 (0.233)	-0.0568 (0.049)	-0.2035 (0.176)
# RS Evals	-0.0000* (0.000)	-0.0004*** (0.000)	-0.0003*** (0.000)	-0.0012*** (0.000)	0.0000 (0.000)	0.0000* (0.000)
<i>Career Characteristics</i>						
Awards (#)	0.0080*** (0.002)	0.0129*** (0.002)	0.0166*** (0.003)	-0.0006 (0.003)	0.0059*** (0.002)	0.0035 (0.003)
AQDs (#)	0.0249*** (0.003)	0.0258*** (0.003)	0.0314*** (0.006)	0.0282*** (0.005)	0.0255*** (0.004)	0.0129 (0.008)
Graduate Educ (0/1)	0.0373*** (0.006)	0.0201** (0.009)	-0.0004 (0.013)	0.0513*** (0.014)	0.0683*** (0.009)	0.0749*** (0.012)
Prior Enlisted	-0.0133 (0.010)	-0.0112 (0.015)	0.0175 (0.025)	-0.0025 (0.024)	0.0189 (0.013)	-0.0598** (0.024)
<i>Demographics</i>						
Female	0.0128* (0.008)	-0.0183 (0.013)	-0.0363** (0.018)	0.0146 (0.017)	0.0278*** (0.009)	0.0233** (0.012)
Black	-0.0217** (0.010)	-0.0456*** (0.018)	-0.0749** (0.031)	-0.0218 (0.021)	-0.0104 (0.012)	0.0078 (0.019)
Non-Black Minority	-0.0292*** (0.007)	-0.0399*** (0.011)	-0.0643*** (0.016)	0.0102 (0.016)	-0.0179* (0.009)	0.0043 (0.013)
Married	0.0442*** (0.007)	0.0228** (0.010)	0.0062 (0.014)	0.0310* (0.017)	0.0592*** (0.010)	0.0571*** (0.014)
Dependents (#)	0.0007 (0.003)	0.0107*** (0.004)	0.0118** (0.005)	0.0039 (0.006)	-0.0056* (0.003)	-0.0076 (0.005)

<i>Past Evaluation Characteristics</i>						
$\Delta(\text{Trait Avg} - \text{RSCA})$ Avg.	0.2774*** (0.014)	0.2174*** (0.019)	0.2324*** (0.027)	0.2347*** (0.036)	0.3728*** (0.022)	0.3024*** (0.036)
EP_Avg.	0.0100 (0.014)	-0.0109 (0.021)	-0.0071 (0.035)	0.0287 (0.033)	0.0425** (0.019)	0.0947*** (0.026)
MP_Avg.	-0.0930*** (0.014)	-0.0827*** (0.020)	-0.1400*** (0.029)	0.0422 (0.033)	-0.0863*** (0.019)	-0.0574** (0.025)
# RS Evals_Avg.	0.0000*** (0.000)	0.0003*** (0.000)	0.0006*** (0.000)	-0.0003*** (0.000)	0.0000*** (0.000)	0.0000** (0.000)
Awards_Avg. (#)	-0.0020 (0.002)	-0.0079*** (0.002)	-0.0100*** (0.003)	-0.0033 (0.003)	-0.0001 (0.002)	-0.0011 (0.004)
AQDs_Avg. (#)	0.0052* (0.003)	0.0058 (0.004)	0.0115* (0.006)	-0.0163*** (0.006)	0.0006 (0.004)	0.0084 (0.010)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
MOS 2 digit fixed effects	Yes	Yes	No	No	Yes	Yes
Board year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Board look fixed effects (up to 4)	Yes	Yes	Yes	Yes	Yes	Yes
Past missing eval indicators	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,711	9,889	5,887	2,520	11,822	4,992
R-squared	0.289	0.328	0.333	0.355	0.292	0.345

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 5: O-4 to O-5 Promotion, Outcome = Promote (0/1)**

	(1) Full Sample	(2) URL	(3) Aviation	(4) Surface Warfare	(5) RL + Staff Corp	(6) Med-Jag
Promote	58%	62%	60%	59%	55%	54%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.1704*** (0.013)	0.1700*** (0.017)	0.2069*** (0.022)	0.1376*** (0.038)	0.1484*** (0.020)	0.2520*** (0.040)
EP	0.2527*** (0.016)	0.2128*** (0.024)	0.1410*** (0.035)	0.1522*** (0.041)	0.2771*** (0.021)	0.2849*** (0.030)
MP	0.0027 (0.016)	-0.0770*** (0.024)	-0.1816*** (0.030)	0.0468 (0.042)	0.0617*** (0.021)	0.0632** (0.029)
1 of 1	0.0602 (0.042)	0.1576*** (0.059)	0.3211*** (0.061)	-0.0607 (0.122)	-0.0602 (0.055)	0.1619 (0.133)
1of1*EP	-0.1717*** (0.043)	-0.2351*** (0.060)	-0.3512*** (0.064)	0.0186 (0.125)	-0.0795 (0.056)	-0.2871** (0.134)
1of1*MP	-0.0244 (0.054)	-0.0022 (0.079)	-0.1077 (0.086)	0.0339 (0.161)	0.0073 (0.069)	-0.2371 (0.154)
# RS Evals	0.0001* (0.000)	0.0002 (0.000)	0.0010*** (0.000)	-0.0000 (0.000)	0.0000 (0.000)	0.0000 (0.000)
<i>Career Characteristics</i>						
Awards (#)	0.0225*** (0.003)	0.0213*** (0.004)	0.0255*** (0.005)	0.0024 (0.009)	0.0202*** (0.004)	0.0194*** (0.005)
AQDs (#)	0.0224*** (0.003)	0.0207*** (0.006)	0.0177** (0.008)	0.0235** (0.010)	0.0218*** (0.004)	0.0106 (0.010)
Graduate Educ (0/1)	0.0580*** (0.009)	0.0287** (0.012)	0.0365** (0.016)	0.0308 (0.028)	0.1064*** (0.015)	0.0839*** (0.020)
Prior Enlisted	-0.0228 (0.045)	0.0837 (0.079)	0.2653** (0.103)	0.1955 (0.177)	-0.0812 (0.055)	-0.1085 (0.091)
<i>Demographics</i>						
Female	0.0141 (0.013)	0.0435 (0.027)	0.0495 (0.038)	0.0626* (0.037)	0.0141 (0.015)	0.0178 (0.019)
Black	-0.0373** (0.015)	-0.0835*** (0.025)	-0.0034 (0.042)	-0.1016*** (0.034)	-0.0175 (0.018)	-0.0167 (0.026)
Non-Black Minority	-0.0408*** (0.011)	-0.0189 (0.018)	0.0138 (0.025)	-0.0342 (0.030)	-0.0574*** (0.015)	-0.0828*** (0.022)
Married	0.0223* (0.013)	0.0189 (0.020)	0.0183 (0.028)	0.0083 (0.036)	0.0237 (0.017)	0.0306 (0.025)
Dependents (#)	0.0033 (0.003)	0.0091* (0.005)	-0.0014 (0.007)	0.0162* (0.009)	-0.0011 (0.005)	-0.0041 (0.007)

<i>Past Evaluation Characteristics</i>						
$\Delta(\text{Trait Avg} - \text{RSCA})$ Avg.	0.3781*** (0.021)	0.3652*** (0.029)	0.3039*** (0.037)	0.4727*** (0.069)	0.3895*** (0.034)	0.4727*** (0.064)
EP_Avg.	0.1009*** (0.019)	0.0561* (0.031)	0.0081 (0.047)	0.0393 (0.053)	0.1273*** (0.024)	0.1979*** (0.038)
MP_Avg.	0.0149 (0.022)	-0.0140 (0.033)	0.0178 (0.044)	-0.0933 (0.062)	0.0248 (0.028)	0.0339 (0.039)
# RS Evals_Avg.	0.0000 (0.000)	-0.0009*** (0.000)	-0.0025*** (0.000)	0.0000 (0.000)	0.0001* (0.000)	0.0001** (0.000)
Awards_Avg. (#)	-0.0102*** (0.003)	-0.0082** (0.004)	-0.0079 (0.005)	-0.0061 (0.009)	-0.0118*** (0.004)	-0.0093* (0.005)
AQDs_Avg. (#)	-0.0008 (0.004)	-0.0040 (0.006)	-0.0083 (0.008)	0.0139 (0.011)	0.0056 (0.005)	0.0019 (0.012)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
MOS 2 digit fixed effects	Yes	Yes	No	No	Yes	Yes
Board year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Board look fixed effects (up to 4)	Yes	Yes	Yes	Yes	Yes	Yes
Past missing eval indicators	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,593	5,046	2,707	1,465	6,547	2,767
R-squared	0.267	0.324	0.369	0.322	0.245	0.307

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 6: O-5 to O-6 Promotion, Outcome = Promote (0/1)**

	(1) Full Sample	(2) URL	(3) Aviation	(4) Surface Warfare	(5) RL + Staff Corp	(6) Med-Jag
Promote	46%	46%	36%	53%	46%	47%
<i>Most Recent Evaluation Characteristics</i>						
$\Delta$ (Trait Avg - RSCA)	0.2541*** (0.014)	0.2528*** (0.020)	0.2199*** (0.025)	0.2526*** (0.045)	0.2604*** (0.020)	0.3196*** (0.042)
EP	0.2006*** (0.018)	0.0570** (0.027)	0.1146*** (0.033)	0.0643 (0.060)	0.3007*** (0.022)	0.3233*** (0.032)
MP	0.0370* (0.021)	-0.0060 (0.032)	0.0340 (0.038)	-0.0156 (0.067)	0.0738*** (0.026)	0.0729** (0.036)
1 of 1	-0.0395 (0.036)	-0.1782*** (0.043)	-0.0482 (0.054)	-0.1463 (0.103)	0.1965** (0.079)	-0.0424 (0.110)
1of1*EP	0.0089 (0.039)	0.1451*** (0.047)	0.0252 (0.061)	0.1181 (0.108)	-0.2059** (0.081)	0.0714 (0.114)
1of1*MP	-0.0434 (0.052)	0.0182 (0.063)	-0.0145 (0.077)	-0.0218 (0.137)	-0.1547 (0.105)	0.2236* (0.136)
# RS Evals	0.0001** (0.000)	-0.0001 (0.000)	0.0002 (0.000)	0.0001 (0.000)	0.0002*** (0.000)	0.0002*** (0.000)
<i>Career Characteristics</i>						
Awards (#)	0.0280*** (0.003)	0.0314*** (0.004)	0.0498*** (0.005)	0.0162** (0.008)	0.0230*** (0.004)	0.0148*** (0.005)
AQDs (#)	0.0264*** (0.002)	0.0400*** (0.003)	0.0422*** (0.004)	0.0486*** (0.007)	0.0181*** (0.003)	0.0091* (0.005)
Graduate Educ (0/1)	0.0677*** (0.016)	0.0267 (0.021)	0.0029 (0.025)	0.1337*** (0.050)	0.0938*** (0.024)	0.0984*** (0.031)
<i>Demographics</i>						
Female	0.0346* (0.018)	0.0300 (0.041)	0.0908 (0.059)	0.0231 (0.057)	0.0273 (0.021)	0.0156 (0.026)
Black	0.0273 (0.022)	-0.0208 (0.041)	-0.0533 (0.072)	0.0380 (0.058)	0.0543** (0.026)	0.0856** (0.043)
Non-Black Minority	-0.0420*** (0.016)	-0.0579** (0.026)	0.0012 (0.040)	-0.0877** (0.042)	-0.0326 (0.020)	-0.0001 (0.028)
Married	0.0466** (0.020)	0.0383 (0.033)	0.0450 (0.043)	-0.0311 (0.060)	0.0513** (0.024)	0.0643* (0.034)
Dependents (#)	0.0048 (0.004)	0.0118* (0.007)	0.0167* (0.009)	0.0066 (0.015)	-0.0006 (0.006)	-0.0005 (0.008)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
MOS 2 digit fixed effects	Yes	Yes	No	No	Yes	Yes

Board year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Board look fixed effects (up to 4)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,368	2,653	1,408	781	3,715	1,687
R-squared	0.266	0.350	0.369	0.365	0.266	0.332

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

<b>Table 7: High and Low Performers, Outcome = Promote (0/1)</b>						
	(1)	(2)	(3)	(4)	(5)	(6)
	O-3 to O-4		O-4 to O-5		O-5 to O-6	
	High Performer	Low Performer	High Performer	Low Performer	High Performer	Low Performer
Promote	76%	41%	68%	16%	52%	12%
<i>Most Recent Evaluation Characteristics</i>						
Δ (Trait Avg - RSCA)	0.1064*** (0.026)	0.1359*** (0.027)	0.1346*** (0.027)	0.0122 (0.030)	0.2353*** (0.031)	0.0716** (0.034)
MP		0.1290*** (0.022)		0.0776** (0.030)		0.0863*** (0.033)
1 of 1	-0.0687*** (0.014)	0.0424 (0.051)	-0.1050*** (0.018)	0.0497 (0.082)	-0.0464* (0.025)	-0.0802 (0.059)
1of1*MP		-0.1200* (0.064)		0.0182 (0.116)		-0.0122 (0.103)
# RS Evals	0.0000 (0.000)	-0.0000 (0.000)	0.0002** (0.000)	0.0000 (0.000)	0.0001 (0.000)	0.0000 (0.000)
<i>Career Characteristics</i>						
Awards (#)	0.0054 (0.004)	0.0086** (0.004)	0.0069 (0.007)	-0.0086 (0.009)	0.0339*** (0.005)	0.0098 (0.008)
AQDs (#)	0.0142** (0.007)	0.0426*** (0.007)	0.0120 (0.008)	0.0365*** (0.013)	0.0269*** (0.004)	0.0365*** (0.009)
Graduate Educ (0/1)	0.0292** (0.014)	0.1087*** (0.020)	0.0767*** (0.020)	0.1074*** (0.030)	0.0058 (0.035)	-0.0458 (0.030)
Prior Enlisted	0.0068 (0.022)	-0.0969*** (0.034)	0.0038 (0.105)	-0.0875 (0.083)		
<i>Demographics</i>						
Female	0.0281 (0.018)	-0.0485** (0.022)	0.0157 (0.027)	0.0518 (0.040)	0.0754** (0.037)	0.0811* (0.049)
Black	-0.0141 (0.021)	-0.0054 (0.030)	-0.0067 (0.031)	0.0288 (0.044)	0.0373 (0.039)	-0.0265 (0.072)
Non-Black Minority	-0.0111 (0.016)	-0.0178 (0.021)	-0.0318 (0.023)	-0.0313 (0.034)	-0.0147 (0.031)	-0.0235 (0.043)
Married	0.0611*** (0.018)	0.0308 (0.021)	0.0072 (0.027)	-0.0145 (0.034)	0.0513 (0.038)	0.0422 (0.046)
Dependents (#)	-0.0071 (0.006)	0.0078 (0.008)	0.0000 (0.007)	-0.0018 (0.011)	0.0134 (0.009)	-0.0000 (0.010)
Past Evaluation Characteristics	Yes	Yes	Yes	Yes	No	No
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
MOS 2 digit fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Board year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Board look fixed effects (up to 4)	Yes	Yes	Yes	Yes	Yes	Yes
Past missing eval indicators	Yes	Yes	Yes	Yes	No	No
Observations	3,800	2,323	2,733	685	1,809	495
R-squared	0.237	0.408	0.233	0.331	0.215	0.292

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Higher performers are defined as individuals who have received an EP in their most recent FITREP and across the past five observed FITREPs. Low performers are defined as individuals who have never received an EP in the last five observed FITREPs.

**Table 8: Enlisted E-6 Sailors, Time to Promote E-5 to E-6 (months)**

	(1) Full Sample	(2) Full Sample	(3) Admin + Deck+Tech +Weapon	(4) Board 2018 or earlier
Time to Promote	<b>52.79</b>	<b>52.79</b>	<b>52.00</b>	<b>51.58</b>
<i>Most Recent Evaluation Characteristics</i>				
$\Delta$ (Trait Avg - RSCA)	-2.1631*** (0.275)		-2.7131*** (0.386)	-2.4165*** (0.365)
Trait Avg		-2.5193*** (0.278)		
RS RSCA		-1.2416*** (0.452)		
EP	-0.9781*** (0.355)	-0.5963* (0.358)	-1.1216** (0.488)	-1.0033** (0.459)
MP	1.0466*** (0.285)	1.2672*** (0.286)	1.0339** (0.402)	1.1516*** (0.362)
1 of 1	3.1613*** (0.535)	3.0522*** (0.535)	3.3237*** (0.686)	3.3005*** (0.681)
1of1*EP	-1.1722** (0.576)	-0.9240 (0.576)	-0.9539 (0.742)	-1.3973* (0.735)
1of1*MP	-0.7598 (0.692)	-0.6693 (0.691)	-0.7405 (0.879)	-0.0006 (0.882)
# RS Evals	-0.0003 (0.001)	0.0003 (0.001)	0.0027*** (0.001)	-0.0011* (0.001)
<i>Career Characteristics</i>				
Awards (#)	-0.7052*** (0.023)	-0.6843*** (0.023)	-0.7631*** (0.032)	-0.6793*** (0.029)
Award Adv Points (#)	0.5158*** (0.026)	0.5133*** (0.026)	0.5218*** (0.035)	0.4270*** (0.032)
Graduate Educ (0/1)	-14.7459*** (0.917)	-14.7732*** (0.917)	-16.6589*** (1.175)	-12.3822*** (1.267)
College Educ (0/1)	-7.9368*** (0.207)	-7.9367*** (0.207)	-8.5628*** (0.276)	-7.5324*** (0.262)
AFQT	-0.2263*** (0.005)	-0.2275*** (0.005)	-0.2633*** (0.007)	-0.2301*** (0.006)
<i>Demographics</i>				
Female	-0.2184 (0.203)	-0.1879 (0.203)	0.2850 (0.272)	0.3184 (0.270)
Black	-1.2932*** (0.213)	-1.1861*** (0.213)	-0.2980 (0.294)	-1.3837*** (0.271)

Non-Black Minority	-0.5303*** (0.166)	-0.4891*** (0.166)	-0.0343 (0.233)	-1.1346*** (0.207)
Married	-1.2807*** (0.164)	-1.2426*** (0.164)	-1.4388*** (0.226)	-1.1639*** (0.211)
Dependents (#)	1.3234*** (0.094)	1.3604*** (0.094)	1.3827*** (0.135)	1.1079*** (0.114)
Age fixed effects	Yes	Yes	Yes	Yes
Indicator for Large Ratings	Yes	Yes	No	Yes
Board year fixed effects	Yes	Yes	Yes	Yes
Observations	64,758	64,758	34,827	36,596
R-squared	0.386	0.387	0.381	0.367

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## APPENDIX TABLES

<b>Appendix Table A1: Robustness Check O-3 to O-4 Promotion, Outcome = Promote (0/1)</b>						
	(1) Full Sample	(2) URL	(3) Aviation	(4) Surface Warfare	(5) RL	(6) Med-Jag
<i>Most Recent Evaluation Characteristics</i>						
Trait Avg.	0.1438*** (0.009)	0.1012*** (0.013)	0.0949*** (0.017)	0.0882*** (0.028)	0.1867*** (0.015)	0.1388*** (0.025)
RS RSCA	-0.1720*** (0.015)	-0.1897*** (0.018)	-0.2306*** (0.027)	-0.1426*** (0.036)	-0.1343*** (0.023)	-0.1558*** (0.034)
EP	0.2198*** (0.012)	0.2209*** (0.019)	0.1983*** (0.026)	0.2692*** (0.040)	0.2034*** (0.016)	0.1800*** (0.021)
MP	0.0733*** (0.012)	0.0947*** (0.018)	0.0269 (0.022)	0.2250*** (0.039)	0.0481*** (0.015)	0.0433** (0.019)
1 of 1	-0.0563** (0.028)	-0.1063*** (0.040)	-0.1330*** (0.043)	0.0941 (0.197)	0.0239 (0.039)	-0.0724 (0.141)
1of1*EP	-0.0405 (0.029)	-0.0065 (0.041)	-0.0063 (0.045)	-0.1517 (0.197)	-0.1089*** (0.039)	0.0267 (0.142)
1of1*MP	-0.0505 (0.036)	-0.0622 (0.050)	0.0130 (0.054)	-0.1072 (0.228)	-0.0588 (0.049)	-0.2141 (0.178)
# RS Evals	-0.0000* (0.000)	-0.0004*** (0.000)	-0.0003*** (0.000)	-0.0012*** (0.000)	0.0000 (0.000)	0.0000 (0.000)
<i>Career Characteristics</i>						
Awards (#)	0.0058*** (0.002)	0.0106*** (0.002)	0.0160*** (0.003)	-0.0008 (0.004)	0.0040* (0.002)	-0.0001 (0.004)
AQDs (#)	0.0221*** (0.003)	0.0227*** (0.003)	0.0303*** (0.006)	0.0278*** (0.005)	0.0227*** (0.004)	0.0095 (0.008)
Graduate Educ (0/1)	0.0384*** (0.006)	0.0239*** (0.009)	0.0028 (0.013)	0.0515*** (0.014)	0.0680*** (0.009)	0.0737*** (0.012)
Prior Enlisted	-0.0139 (0.010)	-0.0104 (0.015)	0.0191 (0.025)	-0.0035 (0.024)	0.0189 (0.013)	-0.0608** (0.024)
<i>Demographics</i>						
Female	0.0138* (0.008)	-0.0129 (0.013)	-0.0312* (0.018)	0.0160 (0.017)	0.0279*** (0.009)	0.0234** (0.012)
Black	-0.0204** (0.010)	-0.0398** (0.018)	-0.0716** (0.031)	-0.0204 (0.021)	-0.0093 (0.012)	0.0074 (0.019)
Non-Black Minority	-0.0285*** (0.007)	-0.0387*** (0.011)	-0.0629*** (0.016)	0.0108 (0.016)	-0.0176* (0.009)	0.0051 (0.013)
Married	0.0439***	0.0225**	0.0069	0.0311*	0.0591***	0.0568***

	(0.007)	(0.010)	(0.014)	(0.017)	(0.010)	(0.014)
Dependents (#)	0.0009	0.0112***	0.0124**	0.0044	-0.0056*	-0.0083*
	(0.003)	(0.004)	(0.005)	(0.006)	(0.003)	(0.005)
<i>Past Evaluation Characteristics</i>						
Trait _ Avg.	0.2525***	0.1938***	0.2444***	0.2352***	0.3517***	0.2439***
	(0.015)	(0.020)	(0.030)	(0.037)	(0.023)	(0.038)
RS RSCA_ Avg.	-0.2824***	-0.2238***	-0.2526***	-0.2393***	-0.3788***	-0.2950***
	(0.014)	(0.019)	(0.027)	(0.036)	(0.022)	(0.036)
EP_ Avg.	0.0626***	0.0525**	0.0025	0.0376	0.0788***	0.1733***
	(0.016)	(0.026)	(0.043)	(0.040)	(0.021)	(0.030)
MP_ Avg.	-0.0498***	-0.0331	-0.1333***	0.0494	-0.0543***	0.0008
	(0.016)	(0.024)	(0.034)	(0.041)	(0.020)	(0.027)
# RS Evals_ Avg.	0.0001***	0.0004***	0.0007***	-0.0003***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awards Avg. (#)	0.0008	-0.0056**	-0.0094***	-0.0030	0.0024	0.0042
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)
AQDs_ Avg. (#)	0.0089***	0.0094**	0.0120*	-0.0155***	0.0042	0.0126
	(0.003)	(0.004)	(0.007)	(0.006)	(0.004)	(0.010)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
MOS 2 digit fixed effects	Yes	Yes	No	No	Yes	Yes
Board year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Board look fixed effects (up to 4)	Yes	Yes	Yes	Yes	Yes	Yes
Past missing eval indicators	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,711	9,889	5,887	2,520	11,822	4,992
R-squared	0.290	0.331	0.336	0.356	0.294	0.349

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Appendix Table A2: O-4 to O-5 Promotion, Outcome = Promote (0/1)**

	(1) Full Sample	(2) URL	(3) Aviation	(4) Surface Warfare	(5) RL	(6) Med-Jag
<i>Most Recent Evaluation Characteristics</i>						
Trait Avg.	0.1813*** (0.013)	0.1849*** (0.018)	0.2409*** (0.025)	0.1519*** (0.038)	0.1554*** (0.021)	0.2522*** (0.040)
RS RSCA	-0.1361*** (0.018)	-0.0983*** (0.024)	-0.0922*** (0.031)	0.0158 (0.062)	-0.1258*** (0.026)	-0.2321*** (0.057)
EP	0.2448*** (0.016)	0.2063*** (0.024)	0.1119*** (0.036)	0.1418*** (0.041)	0.2683*** (0.021)	0.2739*** (0.031)
MP	-0.0026 (0.016)	-0.0845*** (0.024)	-0.1997*** (0.030)	0.0359 (0.042)	0.0565*** (0.022)	0.0594** (0.029)
1 of 1	0.0568 (0.042)	0.1500** (0.059)	0.2933*** (0.064)	-0.0474 (0.125)	-0.0655 (0.055)	0.1571 (0.135)
1of1*EP	-0.1731*** (0.043)	-0.2369*** (0.061)	-0.3439*** (0.066)	0.0015 (0.128)	-0.0769 (0.056)	-0.2846** (0.135)
1of1*MP	-0.0288 (0.054)	-0.0044 (0.079)	-0.1025 (0.087)	0.0022 (0.165)	0.0073 (0.069)	-0.2662* (0.150)
# RS Evals	0.0000 (0.000)	0.0001 (0.000)	0.0009*** (0.000)	-0.0000 (0.000)	0.0000 (0.000)	0.0000 (0.000)
<i>Career Characteristics</i>						
Awards (#)	0.0212*** (0.003)	0.0227*** (0.004)	0.0284*** (0.005)	0.0042 (0.009)	0.0181*** (0.004)	0.0146*** (0.005)
AQDs (#)	0.0202*** (0.004)	0.0224*** (0.006)	0.0207** (0.008)	0.0230** (0.011)	0.0189*** (0.004)	0.0044 (0.010)
Graduate Educ (0/1)	0.0570*** (0.009)	0.0249** (0.012)	0.0329** (0.016)	0.0278 (0.028)	0.1062*** (0.015)	0.0817*** (0.020)
Prior Enlisted	-0.0227 (0.045)	0.0877 (0.078)	0.2653*** (0.098)	0.1847 (0.173)	-0.0822 (0.055)	-0.1110 (0.091)
<i>Demographics</i>						
Female	0.0142 (0.013)	0.0414 (0.027)	0.0491 (0.038)	0.0673* (0.037)	0.0149 (0.015)	0.0189 (0.019)
Black	-0.0378*** (0.015)	-0.0889*** (0.025)	-0.0024 (0.042)	-0.1080*** (0.034)	-0.0168 (0.018)	-0.0146 (0.026)
Non-Black Minority	-0.0408*** (0.011)	-0.0203 (0.018)	0.0144 (0.025)	-0.0395 (0.030)	-0.0567*** (0.015)	-0.0819*** (0.022)
Married	0.0224* (0.013)	0.0153 (0.020)	0.0089 (0.027)	0.0059 (0.036)	0.0248 (0.017)	0.0333 (0.025)
Dependents (#)	0.0033 (0.003)	0.0087* (0.005)	-0.0030 (0.007)	0.0159* (0.009)	-0.0010 (0.005)	-0.0041 (0.007)
<i>Past Evaluation Characteristics</i>						
Trait_Avg.	0.3631*** (0.022)	0.3667*** (0.031)	0.3172*** (0.039)	0.4627*** (0.072)	0.3711*** (0.035)	0.4145*** (0.066)

RS RSCA_ Avg.	-0.3762*** (0.021)	-0.3484*** (0.029)	-0.2766*** (0.038)	-0.4557*** (0.069)	-0.3942*** (0.034)	-0.4722*** (0.064)
EP_ Avg.	0.1214*** (0.022)	0.0304 (0.036)	-0.0532 (0.053)	0.0403 (0.064)	0.1651*** (0.029)	0.2915*** (0.044)
MP_ Avg.	0.0322 (0.024)	-0.0344 (0.036)	-0.0198 (0.046)	-0.0945 (0.068)	0.0577* (0.031)	0.1083** (0.042)
# RS Evals_ Avg.	0.0000 (0.000)	-0.0010*** (0.000)	-0.0026*** (0.000)	-0.0001 (0.000)	0.0001** (0.000)	0.0001*** (0.000)
Awards_ Avg. (#)	-0.0085*** (0.003)	-0.0094** (0.004)	-0.0122** (0.006)	-0.0079 (0.009)	-0.0092** (0.004)	-0.0032 (0.006)
AQDs_ Avg. (#)	0.0017 (0.004)	-0.0069 (0.007)	-0.0125 (0.009)	0.0128 (0.012)	0.0093* (0.005)	0.0102 (0.012)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
MOS 2 digit fixed effects	Yes	Yes	No	No	Yes	Yes
Board year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Board look fixed effects (up to 4)	Yes	Yes	Yes	Yes	Yes	Yes
Past missing eval indicators	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,593	5,046	2,707	1,465	6,547	2,767
R-squared	0.268	0.326	0.376	0.327	0.246	0.311

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

THIS PAGE INTENTIONALLY LEFT BLANK

## LIST OF REFERENCES

- Ahn, S. Y. & Hartmann, L. (2022). *Performance evaluation needs assessment* (NPS-22-N100-A). Naval Postgraduate School.  
<https://apps.dtic.mil/sti/trecms/pdf/AD1201960.pdf>
- Ahn, T., Niven, J., & Veilleux, A. (2021). How long have you been waiting? Explaining the role of irrelevant information in the promotion of U.S. Navy officers. *Economics Bulletin*, 41(2), 604-614.
- Anderson, L., Fryer, R., & Holt, C. (2006). Discrimination: experimental evidence from psychology and economics. In W. M. Rodgers III & W. M. Rodgers (Eds.), *Handbook on the Economics of Discrimination*, (pp. 97–118). Edward Elgar Publishing, Incorporated.
- Baker, B. (2017). *Performance evaluation inflation: Factors affecting the exaggeration of performance ratings in the Air Force*. [Doctoral dissertation, University Honors College, Middle Tennessee State University]. JEWLScholar@MTSU.  
<https://jewlscholar.mtsu.edu/items/a8a99535-7571-43e3-96bf-3aa88d227d76>
- Clemens, A., Malone, L., Phillips, S., & Lee, G. (2012). *An Evaluation of the Fitness Report System for Marine Officers*. Arlington: Center for Naval Analyses.
- Cordial, B. (2017). “Time to Review the FitRep Paradigm.” *Proceedings Magazine*, Vol. 143/12/1378. <https://www.usni.org/magazines/proceedings/2017/december/time-review-fitrep-paradigm>
- Department of the Navy. (2021). *Task Force One Navy Final Report: Our Navy Team – Navigating a Course to True North*.
- Dunst, W. L. (2018). *Evolution of the Marine officer fitness report: a multivariate analysis*. [Master’s thesis, Naval Postgraduate School]. NPS Archive: Calhoun.  
<https://calhoun.nps.edu/handle/10945/58292>
- Ellison, J. C. (2014). *Improving the signal for U.S. Navy officer productivity*. [Master’s thesis, Naval Postgraduate School]. NPS Archive: Calhoun.  
<https://calhoun.nps.edu/handle/10945/44554>
- Evans, L. A. (2018). Simulation-based analysis and optimization of the United States Army performance appraisal system. <https://ir.library.louisville.edu/etd/2906/>

- Evans, L.A. & K. G. Bae. (2019). US Army performance appraisal policy analysis: a simulation optimization problem. *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 16(2): 191-205.
- Evans, L., & Robinson, G. L. (2020). Evaluating our Evaluations: Recognizing and Countering Performance Evaluation Pitfalls. *Military Review*, 100(1), 89.  
[https://digitalcommons.usmalibrary.org/usma\\_research\\_papers/316/](https://digitalcommons.usmalibrary.org/usma_research_papers/316/)
- Hardison, C. M., Zaydman, M., Oluwatola, T., Saavedra, A. R., Bush, T., Peterson, H., & Straus, S. G. (2015). *360-degree assessments: Are they the right tool for the US military?*. RAND National Defense Research Institute, Santa Monica, CA.  
<https://apps.dtic.mil/sti/citations/ADA617217>
- Larger, R. B. (2017). *Effectiveness of the Marine Corps' junior enlisted performance evaluation system. An evaluation of Proficiency and Conduct marks*. [Master's thesis, Naval Postgraduate School]. NPS Archive: Calhoun.  
<https://calhoun.nps.edu/handle/10945/53006>
- Marrone, J. (2020) *Predicting 36-Month Attrition in the U.S. Military: A Comparison Across Service Branches*. RAND Corporation Report.
- Marsh, D. (2020). Evaluating the Evaluation System. *Proceedings Magazine Vol. 146/5/1,407*. <https://www.usni.org/magazines/proceedings/2020/may/evaluating-evaluation-system>
- Small, L. C. (2020). *Successful Practices for Employee Performance Evaluations*. [Master's thesis, Naval Postgraduate School]. NPS Archive: Calhoun.  
<https://calhoun.nps.edu/handle/10945/64880>
- Wolfgeher, S. L. (2009). *Inflation of USAF officer performance reports analyzing the organizational environment*. [Master's thesis, Naval Postgraduate School]. NPS Archive: Calhoun. <https://calhoun.nps.edu/handle/10945/4322>

THIS PAGE INTENTIONALLY LEFT BLANK

## 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