INSTITUTE FOR DEFENSE ANALYSES

Identifying Correlates of Navy Line Officer Retention and Promotion among various Demographic Groups

Machine Learning for Hypothesis Generation

WEAI 2021

Julie Lockwood
Rachel Augustine
Joe King

June 2021
Approved for public release;
distribution is unlimited.
IDA Paper NS P-22655
Log: H 21-000162

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

The authors wish to acknowledge the technical reviews

The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.

About This Publication
This work was conducted by the Institute for Defense Analyses under contract HQ0034-14-D-0001, project CA-6-4854, "Expanding the FIFE for Navy Personnel Management" for the Under Secretary of Defense for Personnel and Readiness. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

For More Information:
Dr. Julie A. Lockwood, Project Leader
jlockwood@ida.org, 703-578-2858

ADM John C. Harvey, Jr., USN (ret) Director, SFRD
jharvey@ida.org, 703-575-4530

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper NS P-22655

Identifying Correlates of Navy Line Officer Retention and Promotion among various Demographic Groups

Machine Learning for Hypothesis Generation

WEAI 2021

Julie Lockwood
Rachel Augustine
Joe King

This page is intentionally blank.

Executive Summary

The Department of the Navy (DoN) desires to better understand the drivers of differences in military officer retention and promotion across demographic groups. To assist efforts to improve representation of females and minorities among Navy senior officers, the DoN asked the Institute for Defense Analyses (IDA) to identify differences in what data features predict retention and O5 promotion outcomes across race/ethnicity and sex groups of Navy officers. The rank of O5 (Commander) is pivotal in many ways: it is the first promotion to require a highly selective promotion board, the first rank to be considered a senior officer, and the first rank with command potential. Attaining the rank of O5 is the gateway to senior Navy ranks. This analysis aims to further the Navy's understanding of factors driving racial/ethnic and sex-based differences in retention and promotion outcomes pertaining to this career milestone.

This analysis identifies features associated with differences in predicted O5 promotion outcomes across race/ethnicity and sex-based groups of Navy line officers commissioned as O1s in 2001-2018. Using administrative military personnel data, we train two machine learning (ML) models: one predicting retention and another predicting promotion. We calculate the effect of each feature on each individual's likelihood of retention or promotion, and compare across demographics to identify differences in which and how much features matter.

We leverage administrative data on military personnel provided by the Defense Manpower Data Center (DMDC) and maintained in IDA's Personally Identifiable Information Enclave (PII Enclave). We analyze restricted and unrestricted Navy officers of the line that commissioned as O1s between 2001 and 2018. We include all regular Active Duty officers, as well as Navy Reservists who have been activated (mobilized) for more than 180 days. Our analytic set comprises 45,006 unique officers who collectively served 338,702 person-years between December 2001 and December 2019. We use these data to construct two tree-based discrete-time machine learning models using IDA's Finite-Interval Forecasting Model (FIFE) version 1.3.4: one model for retention, and a second for promotion to O5. Although FIFE produces retention and promotion forecasts for officers in all years of service and for all future time horizons up to 20 years, this briefing focuses on officers in their tenth year of service in the Navy, with retention and promotion forecast five years into the future.

To avoid immediately attributing differences in retention or promotion probability to race, ethnicity, or sex directly, we do not include information on these demographic characteristics when training the models. One implication of this analytic choice is that to the extent that other features in the data strongly correlate with these excluded demographics, systemic differences in retention or promotion associated with these demographic characteristics may be proxied by other features.

Although beyond the scope of this briefing, follow-on efforts might apply additional analytic tools (currently in a prototype stage at IDA) to identify where relationships discovered by the ML model strongly correlate with various protected class attributes.

To measure the effect of each feature provided to the ML model on an individual's predicted promotion or retention outcome, we use the SHapley Additive exPlanations (SHAP) attribution algorithm. We then calculate and compare feature effects across six demographic groups: White non-Hispanic males, Black non-Hispanic males, Hispanic males (of any race), Other non-Hispanic males (i.e. AIAN, NHPI, mixed-race, and other), White non-Hispanic females, and non-White females. This method illuminates differences across demographic groups in *which* and *how much* features matter for the outcome under consideration. After identifying which features are most consequential for each demographic group, we then assess the degree to which this importance differs across demographics. Because the majority of officers exit military service prior to fulfilling the minimum eligibility requirements for promotion to O5, we examine feature effects from two distinct ML models: one predicting *retention,* and a second predicting *promotion to O5*.

For all demographic groups, we find that many of the most consequential features predicting retention are also the most important predictors of promotion: *officer primary designator*, *officer subspecialty*, *billet designator code*, and *additional officer qualifier designations*. The significance of these career features may intersect with restricted v. unrestricted line status, and requires further investigation. In addition to career features, family and personal attributes (e.g., number of dependents, marriage, citizenship origin, and religious denomination) are highly salient for retention outcomes, while the key features predicting O5 promotions all relate to Navy service regardless of demographic group.

Several important caveats apply to these findings. First and foremost, the relationships we describe are *correlational*, not causal. *Machine learning is a powerful tool that can unearth complex correlations in data, but causality can only be identified when a defensible causal framework exists.* Despite the quality and breadth of the administrative data used in this research, this analysis lacks a causal framework and thus cannot measure or substantiate cause:effect relationships. Absent a causal framework, predictive models like those used here should be viewed as forecast and hypothesis generators. Second, feature effects on the predicted outcome depend on the service year and forecast lead length under consideration.

Moving beyond hypothesis generation and identifying the cause:effect relationships undergirding our results, careful research must identify and exploit experimental or quasi-experimental variation. Many trends identified here are worthy of this level of exploration.

Identifying Correlates of Navy Line Officer Retention and Promotion among various Demographic Groups

Machine Learning for Hypothesis Generation

WEAI 2021

Julie Lockwood
Rachel Augustine
Joe King

July 1, 2021

Navy desires a diverse officer corps

Research questions:

What features most strongly predict mid-career retention and early promotion to O5?

How do these features differ across demographic groups?

Method preview: train Machine Learning (ML) survival models and examine SHAP* values by demographic

Seek to systematically identify complex relationships we might otherwise miss among the millions of possible data interactions

Produce additional descriptives for identified features

Illustrates use of ML for hypothesis generation

Potential element of a disciplined, systematic pre-analysis plan

*SHAP = SHapley Additive exPlanation

Why use ML to address this question?

Regression requires a-priori specification of functional relationship between predictors and output

What relationships might we miss among hundreds of features?

ML can help us identify an optimal functional relationship without requiring the researcher to guess it

Then use other methods to dig deeper…

What features most strongly predict Navy line officer retention through and/or promotion to O5 by YOS* 15?

Occupation features predict retention and promotion:

Primary designator, billet designator, subspecialty, Additional Qualification Designator (AQD) code

Unit Identification Code impacts retention and promotion

Religion, citizenship, marriage, residence type, and duty location predict retention, but not promotion

Does what matters for retention or promotion differ across race-by-sex demographic groups? Yes.

YOS = Year of Service

What is IDA's Finite-Interval Forecasting Engine (FIFE)?

FIFE is a panel data forecasting algorithm and data processor useful for survival analysis and other panel tasks

When applied without a plausibly experimental framework, FIFE is useful for descriptives and forecasting

When applied with a plausibly experimental framework, FIFE can support causal inference just as any other panel estimation routine

FIFE is coded in Python and open-source published at PyPI

We use IDA's research-ready database of Defense Manpower Data Center (DMDC) records

Data source	Description
Active Duty Master	Demographics, career history
Pay	Pay and bonus event data
Deployments	Deployment history
Family/DEERS	Family and dependents

Annual data, December 2001-2019. Total features: 306

Promotion to O5 analysis examines 2001-2004 commissions only

Scope: Active duty Navy officers commissioned as O1s in 2001-2018 as restricted or unrestricted line officers

45k unique officers, 339k person-year observations

Some notes on data and training process

No hold-out set: we are interested in best in-sample prediction, not in out-of-sample prediction

Exclude race/ethnicity and sex from the model

Demographically correlated features will absorb effects of race/ethnicity, gender, and their interactions

Consider five non-overlapping demographic groups when tabulating results

White men – Black men – Hispanic men – Other men

White women – Other women

Proceed in four steps

1. Train models on entire population for each outcome

2. Calculate mean absolute SHAP (SHapley Additive exPlanation) value for each feature within demographic group

3. Identify most important features by demographic group, and differences in importance across groups

4. Observe how category membership differentially influences predictions across demographics

1. Model retention: Who serves long enough to meet O5 board?

Outcome: Exit service vs. remain or right-censored

Population: All individuals in analytic set

Use FIFE's LGBSurvivalModeler to fit a LightGBM binary classifier model for each forecast horizon τ

Obtain probability that each individual remains in service through $t+\tau$, conditional on remaining in prior periods

The cumulative product of predictions from each horizon form an estimated survival function

1. Model promotion: How long until an individual promotes to O5?

Outcome: Promote to O5 vs. remain or right-censored

Population: Individuals in analytic set, if not already exited

Use FIFE's LGBStateModeler to fit a LightGBM binary classifier model for each forecast horizon τ

Obtain probability that each individual achieves O5 rank by $t+\tau$, conditional on remaining in prior periods

2. Calculate mean absolute SHAP value for each feature within demographic group

SHAP (SHapley Additive exPlanation) value for a given feature equals the ***change in the expected model prediction when conditioning on that feature***

Lundberg and Lee (2017) show that Shapley values provide a unified measure of feature importance, and how to estimate them

Lundberg, Scott M., Su-In Lee, "A Unified Approach to Interpreting Model Predictions," *Advances in Neural Information Processing Systems*. 30: 4765–4774. Nov 2017.

3. For the most important features, calculate differences in importance across demographic groups

For each feature, calculate mean absolute SHAP value within each demographic group

We define “important" features as those with SHAP values at or above the 60th percentile

We exclude features that capture period effects (e.g., dates)

For all “important” features, calculate difference in mean absolute SHAP values between each demographic and White men

4. Observe how category membership differentially influences predictions across demographics

For features identified in Step 3, use FIFE's Interacted Fixed Effects modeler to calculate expected 15 YOS outcome

- LHS equals predicted 15 YOS Retention or Promotion status
- RHS contains interacted category and demographic indicators
- Coefficients equal expected 15 YOS Retention or Promotion rates per category X demographic

Excluded small cells

LHS = Left-Hand Side

RHS = Right-Hand Side

Results

Retention from 10 to 15 YOS: Career and personal attributes impact each demographic differently

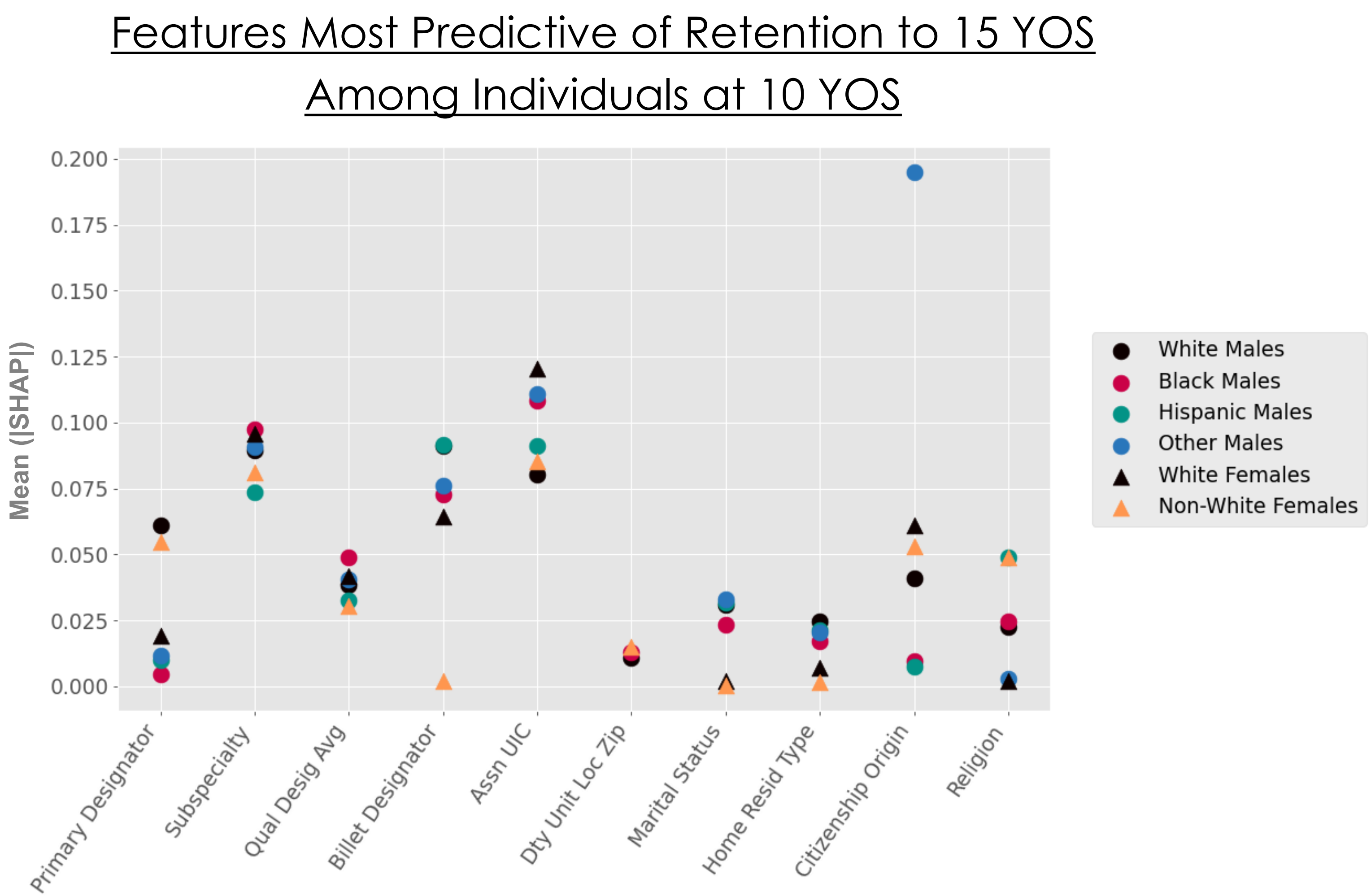

Chart displays the mean absolute SHAP value for each feature, for each demographic group. Feature impact on retention may be positive or negative for any group.

Officer qualifications correlate to differences in retention across demographic groups

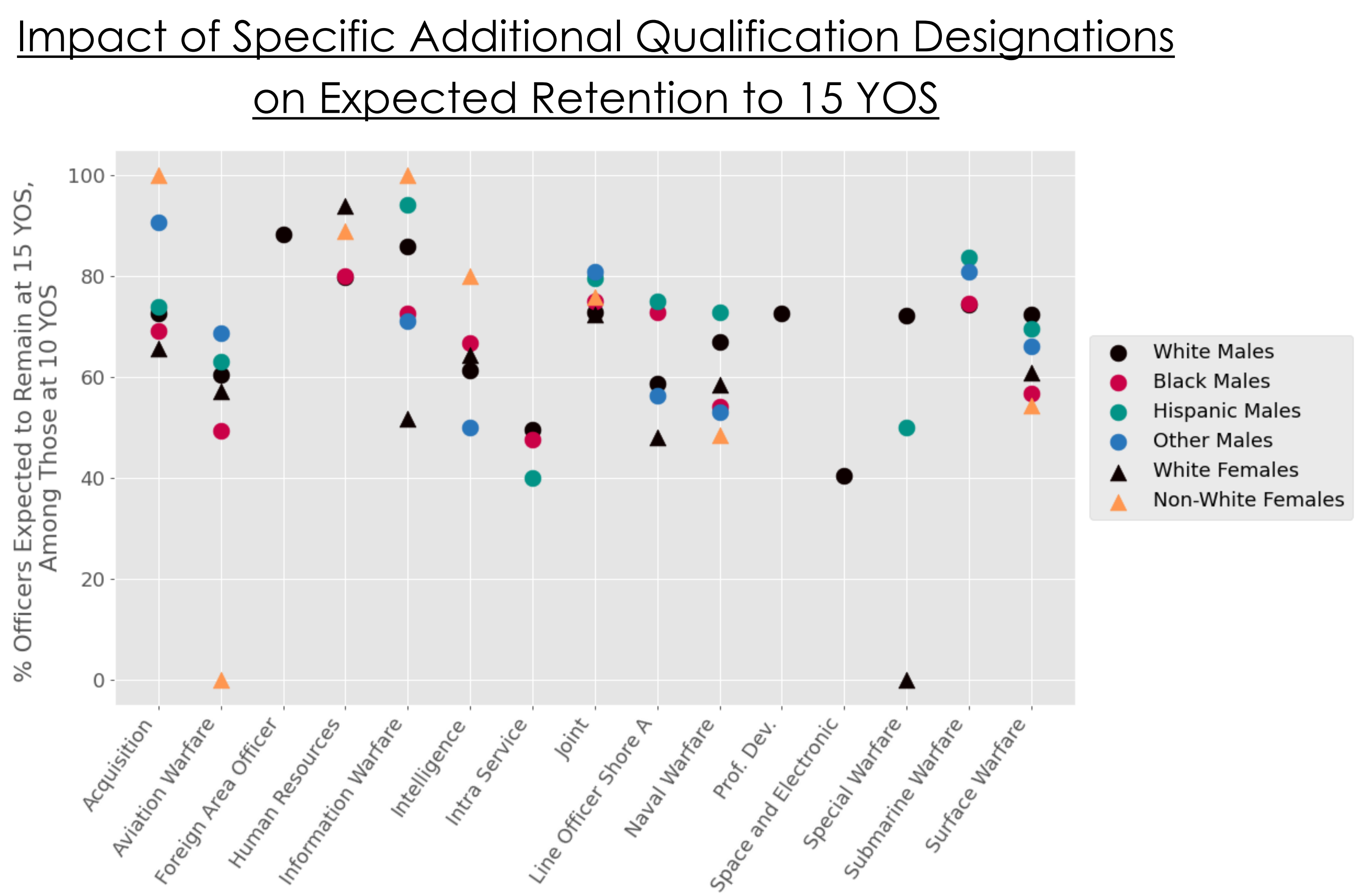

Chart displays the average expected retention of individuals within the indicated category and demographic group. Categories shown contain at least 25 individuals at 10 YOS, with at least 5 in each demographic group at YOS 10, and at least 5 individuals forecast to remain at 15 YOS.

Retention patterns may differ across demographics by religious affiliation

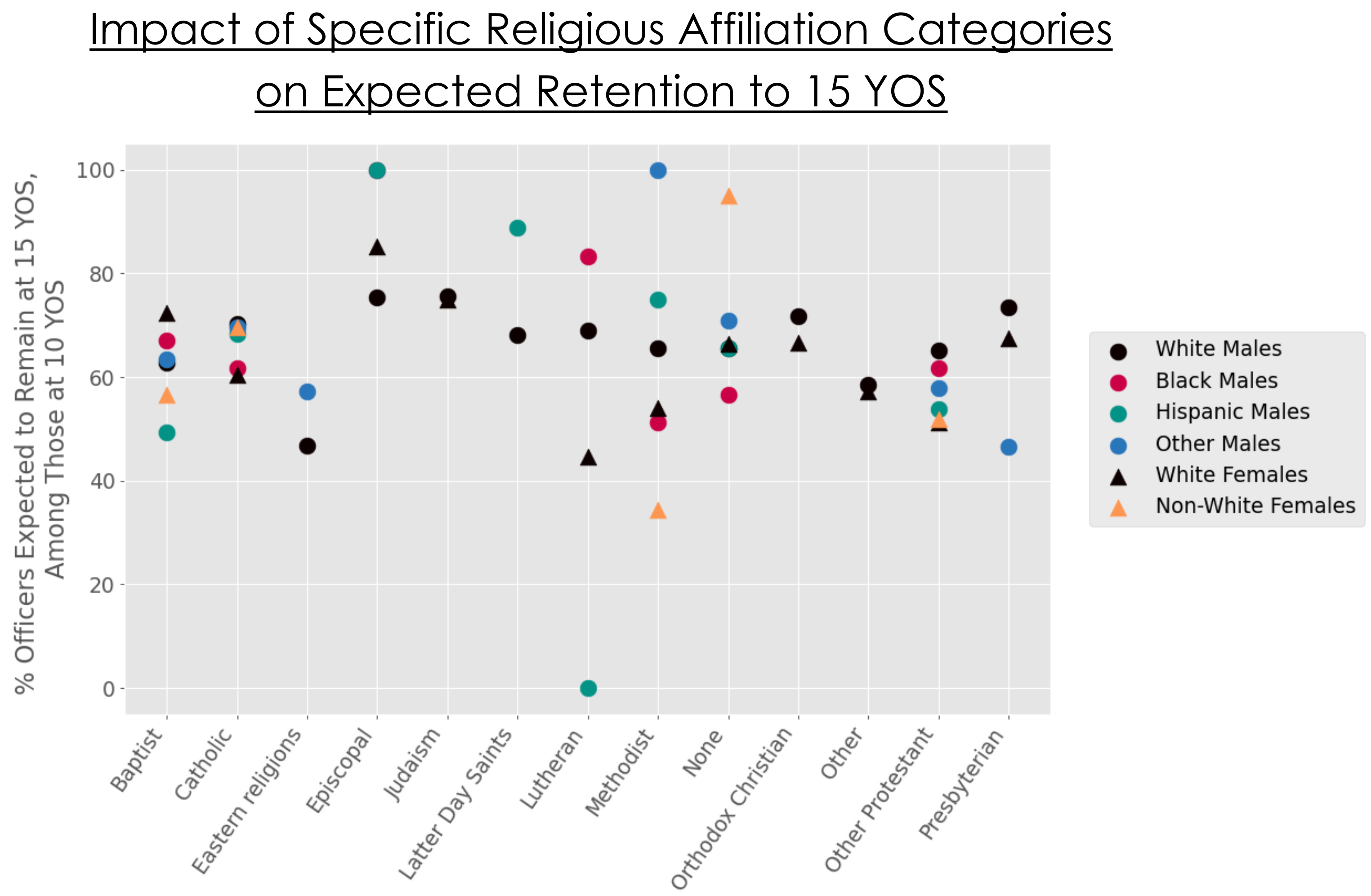

Chart displays the average expected retention of individuals within the indicated category and demographic group. Categories shown contain at least 25 individuals at 10 YOS, with at least 5 in each demographic group at YOS 10, and at least 5 individuals forecast to remain at 15 YOS.

Early promotion to O5 (by 15 YOS): Only career attributes matter; demographic divergences evident

Features Most Predictive of Promotion to O5 by 15 YOS
Among Individuals at 10 YOS

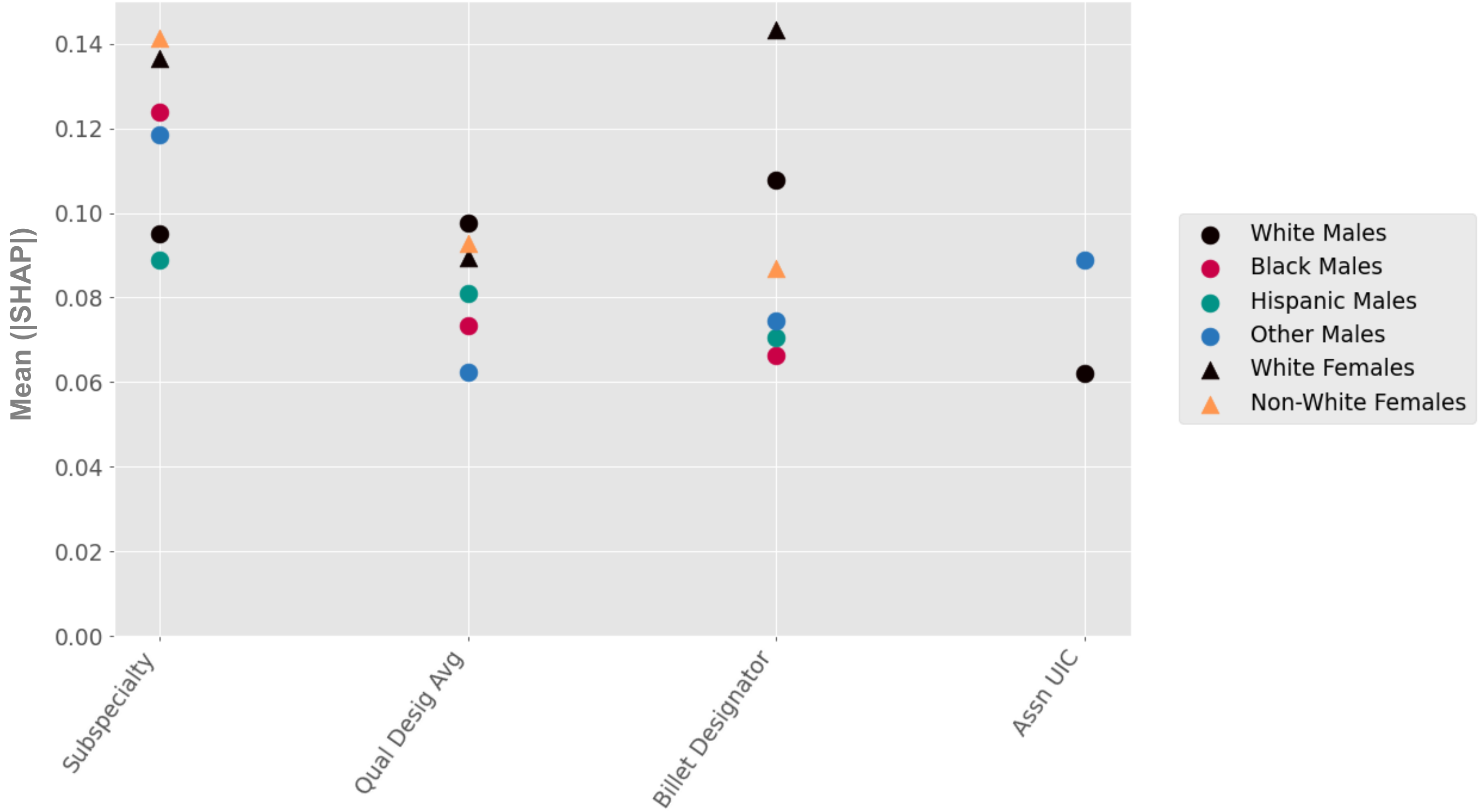

Chart displays the mean absolute SHAP value for each feature, for each demographic group.
Feature impact on promotion may be positive or negative for any group.

Officer qualifications also appear to impact early promotion differently across demographic groups

Impact of Specific Additional Qualification Designations on Expected Promotion to O5 by 15 YOS

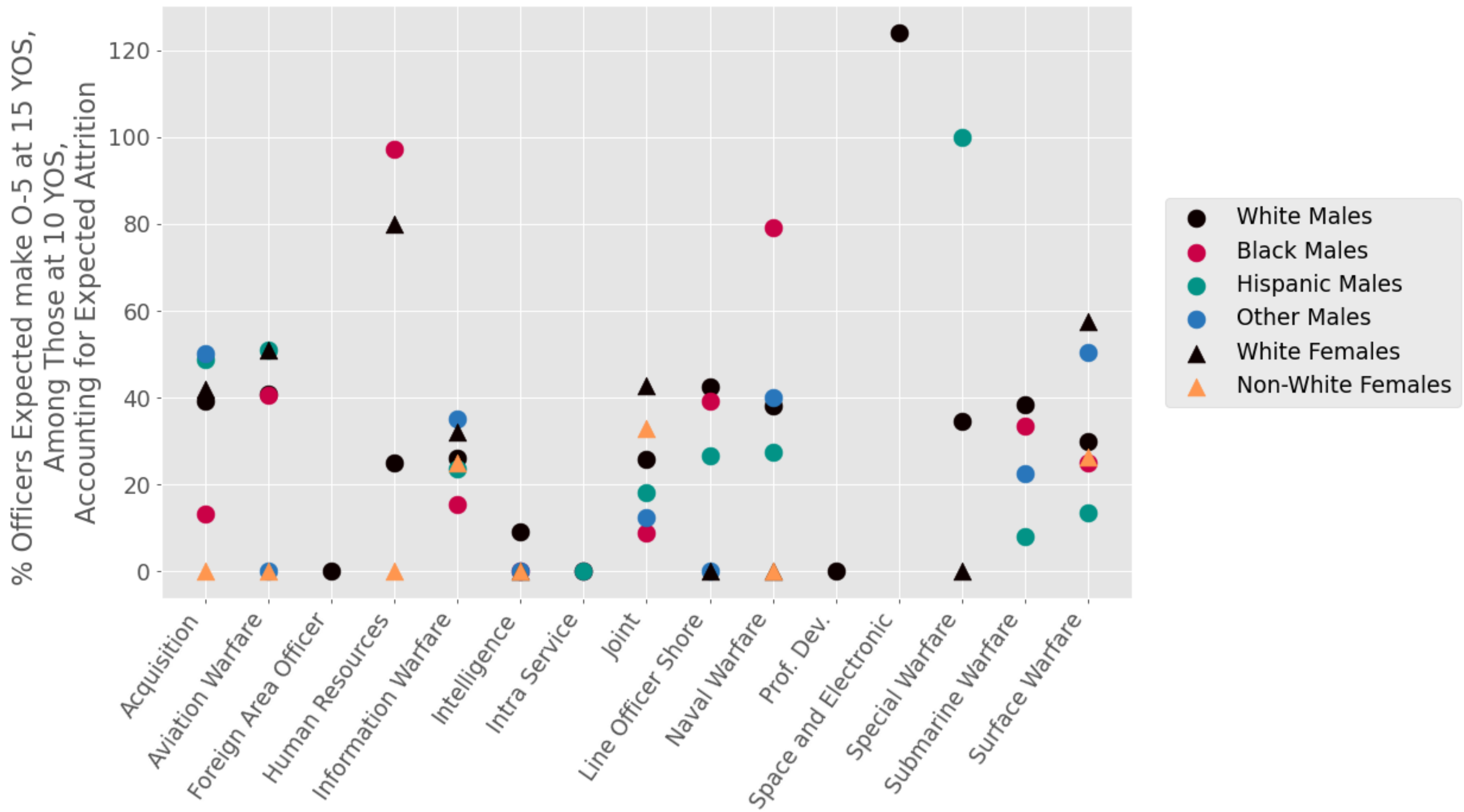

Chart displays the average expected promotion to O5 by 15 YOS of individuals present at 10 YOS within the indicated category and demographic group, adjusted for expected attrition. Categories shown contain at least 25 individuals at 10 YOS, with at least 5 in each demographic group at YOS 10, and at least 5 individuals forecast to remain at 15 YOS.

Exploratory results suggest additional investigation

Why does membership within the same category differentially impact individuals across demographics?

This work suggests many topics for additional exploration…

e.g., UIC, designator, subspecialty, duty location, faith, citizenship

What *doesn't* matter is also interesting

Family factors appear less consequential for females than males

As in any statistical analysis, determining *why* an identified relationship exists requires an experimental framework

Additional research would be needed to confirm any "A causes B" relationships hypothesized by these results

IDA

Appendix

Home residence type codes

Value	Label	Share per value
1	Duty location, with dependents	18.8%
2	Duty location, without dependents	15.8%
3	Residence location, with dependents	1.6%
4	Residence location, without dependents	0.6%
Missing		63.2%

Note: Share per value calculated based on population of all Active Duty service members in DMDC AD Pay file.

Citizenship origin codes

Value	Label	Share per value
A	Born within the U.S., Guam, Puerto Rico, or Virgin Islands	58.4%
B	U.S. citizen, parent became a citizen by naturalization	0.06%
C	Born outside U.S., Guam, Puerto Rico, or Virgin Islands to at least on citizen parent	1.41%
D	U.S. citizen by naturalization	1.71%
Y	Not a U.S. citizen	0.07%
Missing		38.4%

Note: Share per value calculated based on population of all Active Duty service members in DMDC AD Master file.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YY)	2. REPORT TYPE	3. DATES COVERED (From – To)
xx-06-2021	Final	

4. TITLE AND SUBTITLE

Identifying Correlates of Navy Line Officer Retention and Promotion among various Demographic Groups
Machine Learning for Hypothesis Generation
WEAI 2021

5a. CONTRACT NO. HQ0034-14-D-0001

5b. GRANT NO.

5c. PROGRAM ELEMENT NO(S).

6. AUTHOR(S)

Julie Lockwood
Rachel Augustine
Joe King

5d. PROJECT NO.

5e. TASK NO. CA-6-4854

5f. WORK UNIT NO.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882

8. PERFORMING ORGANIZATION REPORT NO.

IDA Paper NS P-22655
Log: H 21-000162

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

OUSD (P&R)
1500 Defense Pentagon, 2E556
Washington, DC 20301

10. SPONSOR'S / MONITOR'S ACRONYM(S) OUSD (P&R)

11. SPONSOR'S / MONITOR'S REPORT NO(S).

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Department of the Navy desires to better understand the drivers of differences in military officer retention and promotion across demographic groups. This analysis identifies features associated with differences in predicted O-5 promotion outcomes across race/ethnicity and sex-based groups of Navy line officers commissioned as O-1s in 2001-2018. Using administrative military personnel data, we train two machine learning (ML) models: one predicting retention and another promotion. We separately calculate the effect of each feature on each individual's likelihood of retention or promotion, and compare across demographic groups to identify differences in which and how much features matter for each group. Findings suggest that retention differences may be structural in nature, and offer several novel insights potentially worth exploring in future research.

15. SUBJECT TERMS

Longitudinal data, panel data, survival analysis, competing risks, multivariate time series, forecasting, machine learning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NO. OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Mr. Lernes Hebert
U	U	U	U	32	19b. TELEPHONE NUMBER (Include Area Code) (703) 697-6631

This page is intentionally blank.