

Lauren Bassett

5/04/2020

Professor Bollinger

ECO491G Spring 2020

Military Recruitment

Introduction

Deciding to join the military is a large commitment made by young adults as they transition to the working world. Young people can choose to enlist after high school, or commission as an officer after college. Military recruiters can be found at high school career fairs, in tents at community events, and on college campuses.

Articles about predatory military recruitment strategies have been published by newspapers from the New York Times to Teen Vogue. The topic yields itself to a discussion about labor economics. People enroll in the military for a variety of reasons: perceived duty, family tradition, professional training, and the extensive benefits. Many enlistments come from the southern regions of the United States. These areas are considered more patriotic, less educated, and of a lower socioeconomic status. Military recruiters, and even the military, are painted as predatory by some. By drawing connections about recruitment, I will be able to either claim that the military is in fact predatory or show that the statistic is not what it looks like on the surface. I want to test the theory that the military preys on poor people, or that confounding variables are at play.

Data Section:

The data in this paper was pulled from the National Longitudinal Survey of Youth (NLS), specifically their study of men and women born between 1980 – 1984, called the NLS Youth 1997, or NLSY97. The NLSY97 follows individuals for 20 years and reports data on many aspects of their lives. To draw conclusions about the military participation rate, I looked at the following variables:

Description of Data (Information that Shapes Data) [*Variable Name in Stata*]

Number of Household Members Working (1997) [*NumOfHHMembersWorking*]

Gross Household Income (1997) [*GrossIncome97*]**Net Worth (1997)** [*NetWorth97*]

Household Members (1997 and 2017) [*HHMembers97*] [*HHMembers17*]

Gender [*Male*] [*Female*]

Age of Participant (1997 – Birth Year) [*AgeIn97*]

Census Region (Northeast, North center, South, West) (2017) [*NE_17*, *NC_17*, *S_17*, *W_17*]

Race (Black, Hispanic, Mixed, Non-Black/Hispanic) [*Black*] [*Hispanic*] [*Mixed*] [*White*]

If the participant moved [*Moved*]

If the mother or father have veteran status [*MotherIsVet*] [*FatherIsVet*]

The portion of their peers who plan on going to college [*AlmostNoPeersGoToCollege*]
[*QuarterOfPeersGoToCollege*] [*HalfOfPeersGoToCollege*] [*MostPeersGoToCollege*]
[*AlmostAllPeersGoToCollege*]

Estimated ASVAB score. [*ASVAB_Estimate*]

Statistics of Interest:

Census Region:

The census regions are categorized by the NSLY as follows:

Northeast states: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont

North Central states: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, North Dakota, South Dakota

Southern States: Alabama, Arkansas, Delaware, Washington DC., Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina

Western States: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming

The four regions used in this study were not balanced by number of states, median incomes, or populations. The northeast region covered 9 states and contained 813 people. The north-central region contained 11 states and 1,154 participants. The southern region covered 12 states and Washington DC and contained 5,464 participants. The west covered 13 states and contained 1,192 participants. One possible limitation of my analysis was the uneven distribution of participants.

Gender:

The gender distribution in this study was almost equal. Out of 5,464 qualifying participants, 2,380 (51.79%) were women and 2,634 (48.21%) were men.

Parental Veteran Status:

Only one participant had a mother who was a veteran, and that individual had no military experience, so the variables relating to Mother's military participation were removed.

Estimated ASVAB Score:

This number is not official data provided by the DoD, but an analysis of the reported ASVAB subtest scores compiled by NLSY to reflect overall ability. ASVAB scores are normally distributed with a mean of 50 and standard deviation of 10¹. To calculate the estimate score, custom sampling weights were assigned based on percentile scores in Mathematical Reasoning, Arithmetic Reasoning, Word Knowledge, and Paragraph Comprehension based on the percentile the participant tested in². The formula

¹ Wall, Janet. 2006. Joining the Military: It Starts with the ASVAB. Accessed 2020. <https://www.military.com/join-armed-forces/asvab/joining-the-military-it-starts-with-the-asvab.html>.

² National Longitudinal Survey of Youth 1997. 2019. ASVAB MATH_VERBAL SCORE PERCENT.

used to create these percentiles is like the Arms Force Qualifying Test (AFQT) score. The acceptance criteria are the same. Those applying for military service must score above a 31 to be considered³.

Gross Income 97:

Taking the unaltered coefficient of Gross Income is not correct in a regression, so to measure the effect of gross income on likelihood of joining the military, we used $\log(\text{GrossIncome97})$.

Model and Estimation approach:

Assumption 1: “The dependent variable consists of two categorical, independent (unrelated groups).”⁴:

Assumption 2: “You have to have two or more independent variables, which should be measured at the continuous or nominal level.”⁵:

Assumption 3: Observations must be independent.

Assumption 4: No multicollinearity

Assumption 5: A linear relationship exists between the independent variables and log odds.

Assumption 6: No outliers, influential points, or leverage points

The analysis was focused on how life conditions affected a person’s likelihood to enlist in the US Military.

Model Summary:

There were 4,137 individuals in my sample. Of these individuals, 103 have military experience.

Estimation Approach:

First, a binomial regression was conducted on MilitaryService and the significant variables were determined. Then I used a combination of stepwise variable selection and VIF scores to determine the best variables to predict the model.

Estimation and testing results:

The simple logistic regression with all models had a reported p-value of 0.000. Thus, the null hypotheses was rejected and our model was statistically significant. We will look at the individual predictors. Below I have included the results as both the odds ratio and logit coefficient.

³ (Wall 2006)

⁴ Laerd Statistics. n.d. Binomial Logistic Regression Analysis using Stata. Accessed 2020.
<https://statistics.laerd.com/stata-tutorials/binomial-logistic-regression-using-stata.php>.

⁵ (Laerd Statistics n.d.)

VARIABLES	(1) Odds ratio	(2) Logit coeff
MilitaryService		
NumOfHHMembersWorking	0.745* (0.129)	-0.295* (0.174)
GrossIncome97	1.000 (5.43e-06)	-5.48e-06 (5.43e-06)
NetWorth_97	1.000 (1.03e-06)	-3.25e-07 (1.03e-06)
HHMembers_97	1.022 (0.0898)	0.0218 (0.0879)
HHMembers_17	1.080 (0.0779)	0.0770 (0.0722)
1.Male	3.085*** (0.829)	1.127*** (0.269)
Agein97	0.900 (0.0799)	-0.106 (0.0888)
1.NE_17	0.0820*** (0.0580)	-2.501*** (0.707)
1.NC_17	0.176*** (0.105)	-1.736*** (0.596)
1.S_17	0.174*** (0.100)	-1.751*** (0.576)
1.W_17	0.199*** (0.115)	-1.613*** (0.577)
1.Black	1.156 (0.427)	0.145 (0.369)
1.Hispanic	2.224** (0.703)	0.799** (0.316)
1.SubMoved	6.230*** (1.527)	1.829*** (0.245)
1.DadServed	3.325 (2.705)	1.201 (0.814)
1.AlmostNoPeersGoToCollege	200,298 (1.700e+08)	12.21 (848.6)
1.QuarterOfPeersGoToCollege	766,524 (6.504e+08)	13.55 (848.6)
1.HalfOfPeersGoToCollege	376,004 (3.191e+08)	12.84 (848.6)
1.MostPeersGoToCollege	430,495 (3.653e+08)	12.97 (848.6)
1.AlmostAllPeersGoToCollege	366,808 (3.113e+08)	12.81 (848.6)
ASVAB_Estimate	1.006 (0.00484)	0.00580 (0.00481)
lnIncome97	1.862** (0.569)	0.621** (0.305)
Constant	3.98e-10 (3.37e-07)	-21.65 (848.6)
Observations	4,098	4,098

seEform in parentheses

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

Coefficients in this model can be interpreted as predicted change in likelihood a participant joins the military. The probability is expressed as $\frac{P(\text{MilitaryService})}{P(\sim \text{MilitaryService})}$

Hosmer-Lemeshow Chi2-test :

Logistic model for MilitaryService, goodness of fit test	Results
Number of Observations:	4098
Number of Groups:	10
Hosmer-Lemeshow chi2(8)	4.16
Prob > chi2	0.8422

Since P was not statistically significant, we failed to reject the null hypothesis. We assumed that the model was a good representation of the data.

Analysis of Statistically significant variables:

In this section, I combine numerical analysis and economic analysis to discuss why the variables may affect the model. I am using the results from the initial regression.

Male:

The odds ratio for men with military experience was 3.18 times that of being a woman. This is not surprising. The military is mostly comprised of men. In 2013, the policy which excluded women from ground combat was lifted.⁶ In 2017, women comprised about 15% of the United States Military.⁷ Women are exempt from the national draft. Enlisting as a female has additional dangers as well. Each year, the Department of Defense (DoD) releases a report on sexual assault within the military. In 2018, they reported that “About 6.2 percent of active duty women indicated experiencing a sexual assault in the year prior to being surveyed.... Using these rates, the Department estimates 20,500 Service members, representing about 13,000 women and 7,500 men, experienced some kind of...sexual assault.”⁸ In comparison, 11 soldiers died that year⁹. There is not as much incentive for women to join the military when other careers are available and have less risk.

⁶ WIIS. 2017. "Women in Ground Combat: Facts and Figures." Women In International Security. February 1. Accessed 2020. <https://www.servicewomen.org/wp-content/uploads/2017/02/Women-in-Ground-Combat-Arms-Fact-Sheet-2-1-17.pdf>.

⁷ (WIIS 2017)

⁸ Inhofe, James M. Fiscal Year 2018. Department of Defense Annual Report on Sexual Assault in the Military. Department of Defense, Washington DC: Department of Defense. https://www.sapr.mil/sites/default/files/DoD_Annual_Report_on_Sexual_Assault_in_the_Military.pdf.

⁹ Bolton, Alexander. 2020. US Combat deaths in Afghanistan highest in years. January 1. Accessed 2020. <https://thehill.com/policy/defense/476461-us-combat-deaths-in-afghanistan-highest-in-years>.

Census Region In 2017:

The odds ratios for the census location of the participant were:

Northeast: 0.082034

North central: 0.176138

South: 0.1736096

West: 0.1992094

At first glance, it seemed odd that every census region decreases odds of enlisting in the military, but this paints an interesting picture of regional service when you compare the odds ratios. The closer the odds ratio is to one, the less likely the predictor will influence the dependent variable. The region with the farthest distance from one is the northeast. Followed by the southern region, then the western region. A participant from the Northeast is less likely to have a Military affiliation than someone from the west. However, the states included in each section are not how I would typically group states into these regions. For example, Maryland, Delaware, and Washington DC are included in the southern region. I wanted to see what was going on, so I used data from the Economic Policy Institute¹⁰ to create a box and whisker plot showing median income. To do this, I took the median income in 2017 of every state and separated the median incomes into the four given regions. Then I constructed a box and whisker plot to highlight the data and show outliers and the general trends.

This helped me understand why my initial hypothesis, that lower income families and therefore more military recruitment, will come from the south. Because the NSYL97 data included high-earning areas like Washington DC were included the Southern region, the average income of the region increases. In this data, there is no way to separate participants by specific state, only by region, so there is no way to account for this outlier within my dataset. This also explains the multicollinearity between log(Income) and region where the participant lives.

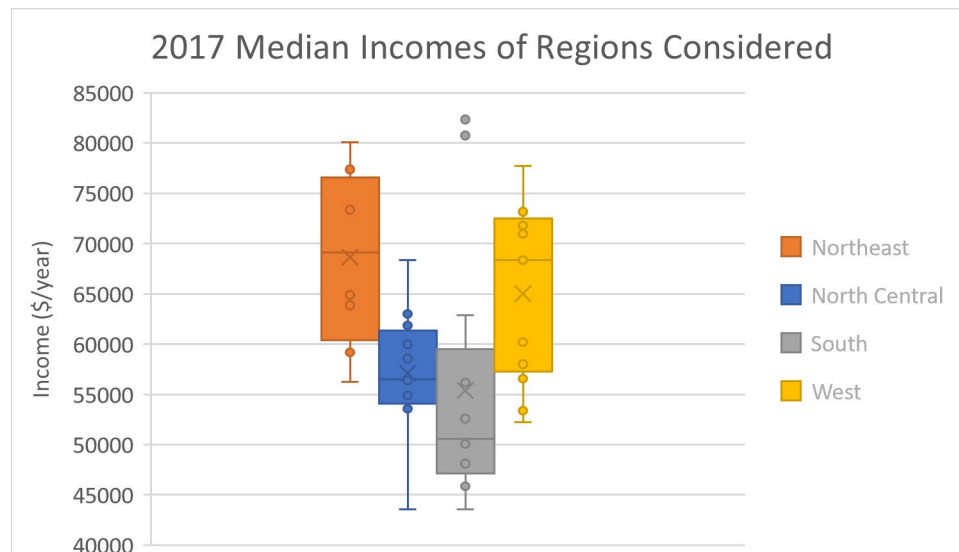


Figure 1: This Box and Whisker Plot shows the distribution of Median incomes across all regions.

¹⁰Julia Wolfe, David Cooper. 2018. "Household incomes in 2017 stayed on existing trends in most states; incomes in 21 states are still below their pre-recession levels." Economic Policy Institute: Working Economics Blog. September 13. Accessed 2020.

Race:

Hispanic:

According to the initial regression, Hispanic individuals in the dataset were twice as likely to have military experience. At first, I thought this was an error; the other race categories were not statistically significant. However, Hispanic Military recruitment is a point of interest for many statisticians. The National Defense Research Institute (NDRI) released a paper observing recruitment rates of Hispanic Americans. Even though Hispanics are a minority group, “[a]mong males, Hispanics had the highest percent responding probably likely or very likely to join relative to white or black males. Among females, 31 percent of Hispanic youth stated they were very likely or probably likely to serve in the military, significantly more than white females”¹¹. The odds ratio for a Hispanic individual is 2.224, which means they are more than twice as likely to be in the Military.

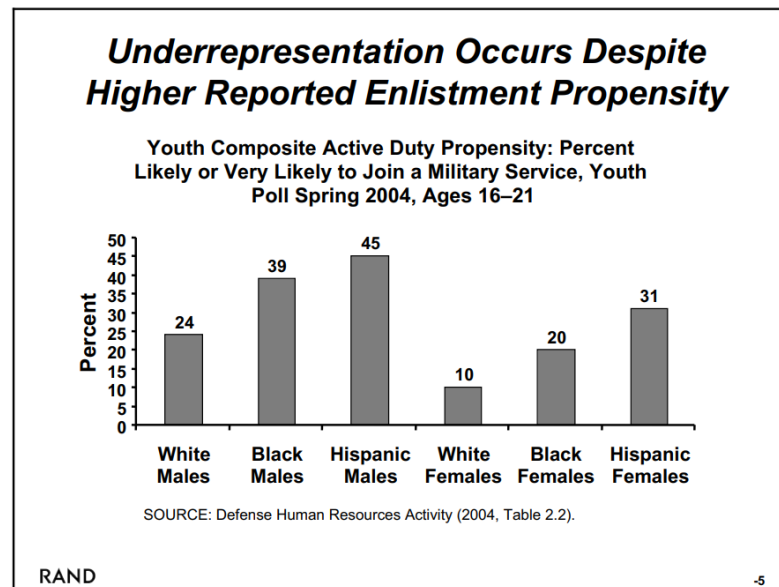


Figure 2: This chart included in NDRI's report shows that for both males and females, the propensity to enlist is higher for Hispanic individuals. (Asch 2005)

Relocating:

Out of the 103 servicemembers represented, 60 had experienced at least one move during the duration of the experiment. Only 43 of the servicemembers had not moved. The odds ratio of moving on Military involvement was 6.230. There are two factors that may contribute to this. The first is that for most who are already enlisted, moving for an assignment is part of the job. If they were active duty during the period from 1997 to 2017, then they are very likely to have moved at least once. The second is that individuals coming from a background where they relocated may not feel tied to a place or career and join the military as a result.

¹¹ Asch, Beth, Christopher Buck, Jacob Alex Klerman, Meredith Kleycamp, and David Loughran. 2005. What Factors Affect the Military Enlistment of Hispanic Youth. RAND National Defense Research Institute.

Log Income:

The log of income was considered statistically significant, with an odds ratio of 1.862. By taking the log value of income, the value was converted to a linear value. In this example, a 1% increase in income increased the likelihood of the individual enlisting in the military. This had the opposite effect that I considered in my hypothesis. In further testing, it became clear that this variable had collinearity with many other variables, specifically, census region.

ASVAB scores:

Predicted ASVAB Scores were not statistically significant in determining likeliness of joining the military (even if they were, the reported odds ratio is 1.006). However, because typically ASVAB scores are used to determine a persons' ability before they join the army, I found it important to discuss why the data was not significant. To do this, I compared the scores of the participants who joined the military to the whole sample. The mean ASVAB score of the participants in the military are shown below.

ASVAB Scores for Participants With Military Service			
Mean Estimation	Mean	Std Error	[95% Confidence Interval]
ASVAB_Estimate	51.174566	2.598307	46,59193, 56.89939

The mean and standard deviation were relatively high. This should have indicated that high scores correlate with military enlistment. However, look this was not the case compared to the mean for all participants.

ASVAB Scores for all Participants			
Mean Estimation	Mean	Std Error	[95% Confidence Interval]
ASVAB_Estimate	45.68577	0.3973573	44.90679, 46.475

They are lower, but it is important to look at the entire picture. The minimum required ASVAB score to join the military is 31¹². The average participant scored above that, and so the ASVAB scores were not a good indicator on who would enlist or not. Because of this, the ASVAB score did not affect likeliness of joining the military in this model.

¹² (Wall 2006)

Additional Analysis:

After seeing which variables were statistically significant in the full model, I generated another binomial logit regression on only the statistically significant variables.

VARIABLES	(1) odds ratio	(2) Coefficient
MilitaryService		
Male	3.122*** (0.821)	1.138*** (0.263)
NE_17	0.0696*** (0.0463)	-2.665*** (0.665)
NC_17	0.129*** (0.0713)	-2.048*** (0.553)
S_17	0.136*** (0.0718)	-1.992*** (0.526)
W_17	0.169*** (0.0901)	-1.778*** (0.533)
Hispanic	1.829** (0.512)	0.604** (0.280)
lnIncome97	1.354** (0.186)	0.303** (0.137)
Constant	0.00241*** (0.00380)	-6.028*** (1.577)
Observations	4,137	4,137

seEform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All the data was statistically significant, but income still increased likelihood of military service. Next, I checked the VIF scores for the variables in this regression.

The values of the VIF scores below were too high to be considered. The other variables are not depicted in the table below because their VIF score was below 2.0.

Variable	VIF(1)	1/VIF (2)
lnIncome97	51.05	0.019591
S_17	21.25	0.047062
NC_17	13.01	0.076877
W_17	12.68	0.078863
NE_17	7.68	0.130126

I then tested the three variables of interest for multicollinearity: log Income, Gender, and Census Location all are statistically significant.

Military on lnIncome		Military on Male		Military on Location	
Adjusted LR	4.17	Adjusted LR	31.93	Adjusted LR	18267.83
Prob > chi2	0.000000	Prob > chi2	0.0000	Prob > chi2	0.000000

In order to control the multicollinearity, I tested the model without lnIncome and without location.

First: removing log (Income)

In the model without log(Income), being Hispanic was no longer statistically significant. The column on the right displays the odds ratios with the Hispanic indicator, and the column on the left shows the odds ratios without the Hispanic indicator.

VARIABLES	(1) odds ratio	VARIABLES	(1) odds ratio
MilitaryService		MilitaryService	
Male	3.407*** (0.791)	Male	3.402*** (0.790)
NE_17	0.0690*** (0.0386)	NE_17	0.0672*** (0.0375)
NC_17	0.129*** (0.0616)	NC_17	0.125*** (0.0589)
S_17	0.151*** (0.0665)	S_17	0.147*** (0.0645)
W_17	0.208*** (0.0929)	W_17	0.209*** (0.0933)
Hispanic	1.155 (0.281)	x	x
Constant	0.0559*** (0.0254)	Constant	0.0588*** (0.0262)
Observations	5,464	Observations	

seEform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In the new model excluding Hispanic indicators, all variables were considered statistically significant. Comparing odds ratios of the northeast, north center, south, and west, a person from the West is more likely to enlist, followed by a person from the south, then the central northern area. An individual from the north was the least likely to enlist.

The result of the VIF Scores for this model, excluding the Hispanic indicator are as follows:

Variable	VIF(1)	1/VIF (2)
Male	1.91	0.524142
S_17	1.35	0.742414
NC_17	1.21	0.823939
W_17	1.20	0.833571
NE_17	1.15	0.871397

The VIF scores for this regression are satisfactory, however, in the original model, the location dummy variables also had high VIF scores, so I want to check and see how removing them and reinserting log(income) affects the model.

Second: removing location data

A similar logit regression was performed using the significant variables, except for the location dummy variables.

VARIABLES	(1) odds ratio
Male	3.096*** (0.812)
Hispanic	2.044*** (0.545)
lnIncome97	1.377** (0.188)
Constant	0.000280*** (0.000414)
Observations	4,137

seEform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The result of the VIF Scores for the new model are:

Variable	VIF(1)	1/VIF (2)
lnIncome97	2.10	0.476254
Male	1.92	0.520911
Hispanic	1.19	0.837620

Since acceptable VIF scores were below 2.00, this interpretation of the data was not satisfactory.

Conclusions:

Conclusions:

Based on my analysis, there were factors that contributed to an individual's likelihood that they enlisted in the Military. While I was unable to find a statistically significant relationship between Military Enrollment in 2017 and the log of that individual's Gross Income in 1997 that was not affected by other variables. However, I believe some of this was due to how the individuals in the sample were grouped together. The regional groups, especially the south, have large outliers in Gross Income because Washington DC and Maryland are considered in that region. When log income and region were included, there was multicollinearity because of changes in income in each state. If the state groupings were closer together in terms of socioeconomic status, I believe the results would have been clearer.

There were some findings in my data that genuinely surprised me. Hispanic men and women were more likely to enlist than any other race group. I was also surprised to see how much more likely men are to enlist than women in the military. One of the things that I did not include in my research was the statistics of sexual assault outside of the military. It is possible that the rate of women who are sexually assaulted within the military is comparable to the rate in general.

Moving forward, I would be interested to see how the individual states effect military recruitment instead of larger regions. I would also look more specifically at socioeconomic status. It may have been more valuable to look at a family's gross income the year the participant turned 18 instead of gross income in 1997 for all participants.

Bibliography

- Asch, Beth, Christopher Buck, Jacob Alex Klerman, Meredith Kleycamp, and David Loughran. 2005. *What Factors Affect the Military Enlistment of Hispanic Youth*. RAND National Defense Research Institute.
- Bolton, Alexander. 2020. *US Combat deaths in Afganistan highest in years*. January 1. Accessed 2020. <https://thehill.com/policy/defense/476461-us-combat-deaths-in-afghanistan-highest-in-years>.
- Inhofe, James M. Fiscal Year 2018. *Department of Defense Annual Report on Sexual Assault in the Military*. Department of Defense, Washington DC: Department of Defense. https://www.sapr.mil/sites/default/files/DoD_Annual_Report_on_Sexual_Assault_in_the_Military.pdf.
- Julia Wolfe, David Cooper. 2018. "Household incomes in 2017 stayed on existing trends in most states; incomes in 21 states are still below their pre-recession levels." *Economic Policy Institute: Working Economics Blog*. September 13. Accessed 2020.
- Laerd Statistics. n.d. *Binomial Logistic Regression Analysis using Stata*. Accessed 2020. <https://statistics.laerd.com/stata-tutorials/binomial-logistic-regression-using-stata.php>.
- National Longitudinal Survey of Youth 1997. 2019. *ASVAB MATH_VERBAL SCORE PERCENT*.
- Wall, Janet. 2006. *Joining the Military: It Starts with the ASVAB*. Accessed 2020. <https://www.military.com/join-armed-forces/asvab/joining-the-military-it-starts-with-the-asvab.html>.
- WIIS. 2017. "Women in Ground Combat: Facts and Figures." *Women In International Security*. February 1. Accessed 2020. <https://www.servicewomen.org/wp-content/uploads/2017/02/Women-in-Ground-Combat-Arms-Fact-Sheet-2-1-17.pdf>.