BETH ANN GRIFFIN, GEOFFREY E. GRIMM, ROSANNA SMART,
RAJEEV RAMCHAND, LISA H. JAYCOX, LYNSAY AYER, ERIN N. LEIDY,
STEVEN DAVENPORT, TERRY L. SCHELL, ANDREW R. MORRAL

Comparing the Army's Suicide Rate to the General U.S. Population

Identifying Suitable Characteristics, Data Sources, and Analytic Approaches

For more information on this publication, visit www.rand.org/t/RR3025

Library of Congress Cataloging-in-Publication Data is available for this publication.
ISBN: 978-1-9774-0359-9

Published by the RAND Corporation, Santa Monica, Calif.

Preface

This report documents research and analysis conducted as part of a project entitled *High-Risk Civilian Population Suicidality*, sponsored by the Office of the Deputy Chief of Staff, G-1, U.S. Army. The purpose of the project was to identify relevant civilian databases that would be suitable for creating comparison proxies for the Army to better compare its suicide rate, as well as provide validated information on risk and protective factors for suicide in civilian settings that may be common to Army soldiers. Also, the project sought to analyze the degree to which suicide rates and possibly other measures of suicidality (ideation, attempts) differ between the Army and the civilian comparison population.

This research was conducted within RAND Arroyo Center's Personnel, Training, and Health Program. RAND Arroyo Center, part of the RAND Corporation, is a federally funded research and development center sponsored by the U.S. Army.

RAND operates under a "Federal-Wide Assurance" (FWA00003425) and complies with the *Code of Federal Regulations for the Protection of Human Subjects Under United States Law* (45 C.F.R. 46), also known as "the Common Rule," as well as with the implementation guidance set forth in Department of Defense (DoD) Instruction 3216.02. As applicable, this compliance includes reviews and approvals by RAND's Institutional Review Board (the Human Subjects Protection Committee) and by the U.S. Army. The views of sources utilized in this study are solely their own and do not represent the official policy or position of DoD or the U.S. government.

Contents

Figures and Tables

Figures

Tables

Summary

Over the past 15 years, the suicide rate among members of the U.S. armed forces has doubled, with the greatest increase observed among soldiers in the Army (Mancha et al., 2014). This increasing rate is paralleled by a smaller increase the general U.S. population (Curtin, Warner, and Hedegaard, 2016), observed across both genders, in virtually every age group, and in nearly every state. An empirical question exists: What is the extent or degree to which the suicide trend in the Army is unique to the Army, relative to what is observed in the general population?

The Army has typically attempted to address this question by standardizing the general population to look like the military population on demographic characteristics (e.g., age and gender are used most frequently [Watkins et al., 2018; Reimann and Mazuchowski, 2018] and race/ethnicity as well on occasion) (Ramchand et al., 2011). Standardization aims to make the general population look like the Army population on the characteristics being used in the procedure, thereby allowing for comparisons that are done on populations with the same characteristics and minimizing the ability of the included characteristics to explain the observed rate differences.

However, given the rise in suicide rates over the past decade, the Army wanted to better understand whether standardization based solely on age and gender is enough. Expanding the characteristics on which the general population is standardized to match the Army could be useful to gain a better understanding of the suicide trends in the Army. However, such a change also brings with it some challenges. First, changing the characteristics that are included in the standardization inherently changes the underlying suicide rate for the general U.S. population since there is a shift in the type of Army population with which the matched general population is being compared. In addition, expansion of the characteristics still results in having a large number of unmeasured factors that cannot be included in this type of analysis.

In this report, RAND Arroyo Center investigated how accounting for additional population risk factors beyond age and gender affects suicide rate differences between soldiers and a comparable subset of the general U.S. population. This is a technical report that surveys data and methods available to improve how the suicide rates of the two populations are being compared. The report does not aim to estimate the causal effects of Army service on suicide.

Conceptually, there are three different types of factors on which we might aim to match or standardize the general U.S. population to look like Army.

Set I includes demographic characteristics, such as age, gender/sex, race/ethnicity, and geography of origin. These factors represent stable characteristics that differ between the Army and the general U.S. population because those who choose (or are allowed) to enter the Army are not fully representative of the general population.

Set II includes characteristics that might be somewhat influenced by Army service, such as marital status and education. These factors are largely determined by individual service members' personal interests and aptitudes that generally predate their military service but may also reflect military policy or opportunities that occurred as part of military service.

Set III includes characteristics that are known to be directly affected by military policy or experiences, such as access to firearms, mental health, or occupation.

Standardization on the first set is a minimum requirement for identifying the effect of Army service on suicide risk. However, most research to date considers only age and gender when comparing Army suicide risk with the general population (Watkins et al., 2018; Reimann and Mazuchowski, 2018). If the goal of the comparison is to determine whether suicide risk is higher or lower for soldiers than it would have been if they had not been in the Army, it is critical to control for the fact that individuals who chose to join the Army look substantially different on these core demographic differences from the general population, even before they joined. However, even matching on a full set of demographic variables could be misleading given the number of unmeasured factors on which Army servicemembers differ from the general U.S. population. For example, the kinds of people who join the Army may be more "psychologically resilient" than individuals in the general U.S. population but may face much higher stresses than individuals in the general population, resulting in the adjusted suicide rates (for the measured covariates we standardized on) being the same for soldiers as the general U.S. population even though the suicide rate of soldiers would have been lower than for individuals in the general population in the absence of the unadjusted special stressors to which Army personnel are exposed.

In contrast, controlling for the characteristics in the third set needs to be carefully considered because matching on these types of characteristics would dramatically alter how one interprets any differences between the Army and the matched general population suicide rates. For example, if the goal of an analysis is to estimate the effect of Army service on suicide risk, including such controls as access to firearms and mental health should be avoided. This is because these types of characteristics represent the specific mechanism by which Army service could affect suicide risk (e.g., through availability of military firearms or mental health problems due to military trauma). An analysis that would attempt to match on these types of factors would obscure the fact that suicide risk is higher or lower for soldiers specifically because they joined the Army and gained access to these mechanisms. Any differences in suicide rates between the Army and matched general population group in these types of analyses should therefore be interpreted cautiously. However, there may be some research purposes for which one does want to control for such factors. For example, if one wants to examine suicide differences between the Army and the general population for a subgroup of individuals with specific mental health diagnoses or to understand how much of the effect of Army service on suicide risk is mediated through access to military firearms, then it may be important to match the populations on these factors.

Characteristics in the second set also warrant careful consideration before inclusion in the standardization process. The decision to include factors in Set II that might be somewhat influenced by Army service, such as marital status and education, depends primarily on the underlying assumptions about the relationship between Army service and these factors. For example, soldiers are more likely to be married than similarly aged individuals in the general population. One potential explanation for this difference is that the Army attracts individuals who also have an interest in getting married (perhaps they are more religious or socially conservative than the general population, or they have a greater interest in having children). If this is the correct theory, marital status should be included in the matched comparisons between the Army and the general population in the same way as variables in Set I. Alternatively, if the higher marriage rate among soldiers reflects Department of Defense (DoD) policies designed to promote marriage under the belief that marriage makes for healthier soldiers, then matching on marital status will produce a comparison between Army and the general population that ignores the effect of Army policies designed to promote soldier well-being. Under this second type of theory, one should avoid matching on this factor when trying to estimate the effect of Army service on suicide risk in the same way that they should be careful with variables in the third set of factors. In short, there are some characteristics for which it is not completely clear whether they should or should not be controlled for in any comparison between soldiers and the general population. Those decisions will need to be informed by the researchers' and Army's theory about the relationship of such factors to Army service and the specific goals of the analysis.

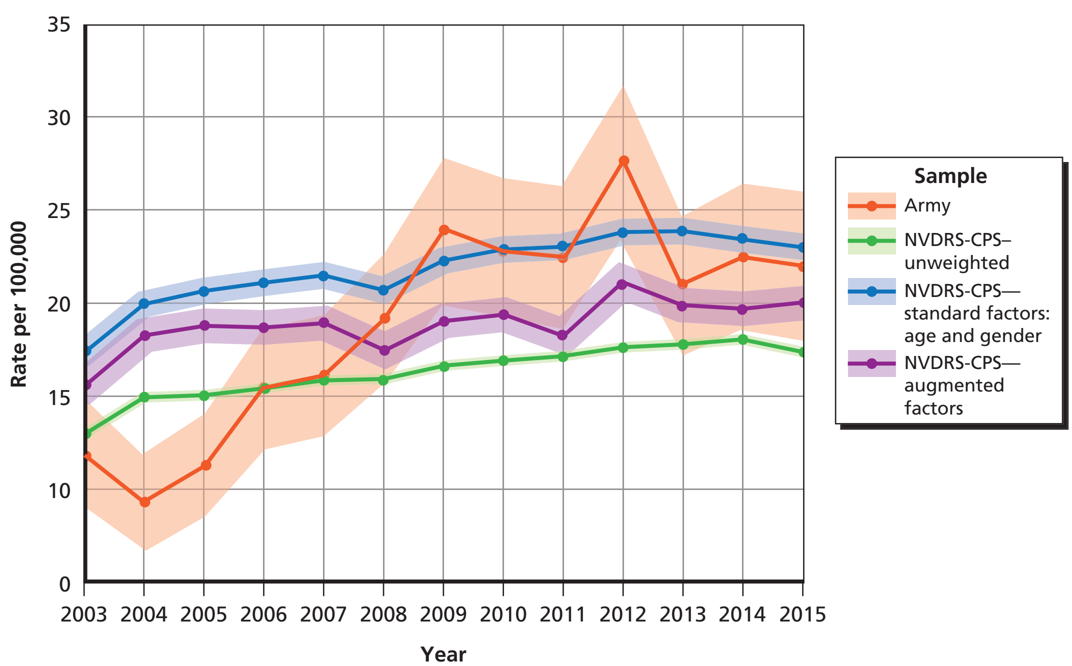
In this report, we explore the various characteristics included in these three sets in more detail in Chapter Two of this report. For many characteristics of interest, data are lacking that would allow for a full exploration of the implications of matching on factors that are included in Sets II and III. Nonetheless, we identified six factors that are related to suicide in the general population and/or the Army, that differ in frequency between the two populations, and that have data available for comparing the Army and general population suicide population. These "matchable factors" are gender, age, time, race/ethnicity, marital status, and educational attainment. We explored the impact of including each factor in comparisons that standardize a subset of the general population to look like the Army as well as the conceptual implications of different sets of "matchable" factors.

Several databases are available to examine suicide risk in the general population (described in Appendix C). Our goal was to select one that was representative of the U.S. population, or a subset of it, and that included an expanded set of factors with which to match to the Army sample. In terms of suicides themselves, the National Violent Death Reporting System (NVDRS) is the only available state-based reporting system that pools data from multiple sources into a usable, publicly available database on violent deaths. However, the NVDRS contains only detailed data on suicides. To establish whether marital status, education level, or other key factors are associated with suicide risk among the general U.S. population, we also needed information on whether those characteristics are over- or underrepresented among the suicide cases as compared with the more general population from which the suicide cases arose. Therefore, we merged the suicide cases from the NVDRS with the Current Population Survey (CPS), a general population database that contains representative data of the general populations living in the states included in the NVDRS in each year. The CPS is a nationally representative and state-by-state representative survey providing high-quality information about the characteristics of the general U.S. population overall and the population of the NVDRS states.

A current limitation of the NVDRS is that it contains only suicide information on a subset of states during our study period. Currently the NVDRS contains data on just 27 states, excluding some of the most populous states in the country. Despite this limitation, we opted to use the NVDRS because of the rich covariate information given on each suicide case. Future plans include expanding the NVDRS to 40 states that will allow future work replicating our methods to have greater generalizability than the results we report. We note that we subset the CPS to only those states included in the NVDRS in each year. Thus, the combined data sets (our NVDRS-CPS sample) provide us with usable information on marriage and educational categories for those who died by suicide and for those who did not, a necessary condition for estimating suicide risk. Because geographic differences exist in state suicide rates, suicides in the NVDRS-CPS sample are likely to differ in systematic ways from suicides nationally, so the NVDRS-CPD sample used in this report is not representative of the entire general U.S. population. As a result, a comparison of Army rates to the NVDRS-CPS sample drawn from just the states participating in the NVDRS may not be used to understand how risk in the Army differs from the risk typically experienced nationally. We note that the subset of NVDRS states have slightly higher suicide rates than the general U.S. population, as presented in Chapter Four. Nonetheless, we believe the lessons learned on different sets of matching factors using the NVDRS-CPS sample provide meaningful information for future efforts that match the Army to the general population since NVDRS will be expanding to more states and both NVDRS and CPS offer rich data sources for future comparisons between the Army and the general U.S. population.

Figure S.1 illustrates how using different factors to standardize our NVDRS-CPS sample to the Army in a given calendar year affects the implications one might draw from comparing Army suicide rates with the general U.S. population. First, as is well known, both populations have experienced an increase in suicide rates since 2003, with the scale of the increase being larger in the Army. When adjusting for only age and gender, the Army suicide rate is significantly lower than the NVDRS-CPS suicide rate before 2008. After 2008, the confidence bands for these two curves (Army versus NVDRS-CPS, adjusted for gender and age) are largely overlapping, suggesting little difference between the two populations within each year.

Figure S.1
Army and Civilian Suicide Rates for Unweighted NVDRS-CPS, NVDRS-CPS Weighted Using Standard Factors (Age Plus Gender), and NVDRS-CPS Weighted Using Augmented Factors (Age, Gender, Race/Ethnicity, Marital Status, and Educational Attainment)



When we expand the standardization to include race/ethnicity, education, and marital status, the "expected" suicide rate in the weighted NVDRS-CPS sample is consistently lower in each calendar year than in the NVDRS-CPS curve that used only age and gender in the adjustment. This conceptually makes sense since we have expanded the characteristics on which we want to make our NVDRS-CPS sample look like the Army, including characteristics that might themselves be impacted by military service like education and marriage. In this fully adjusted analysis, the Army again has significantly lower rates in 2003, 2004, and 2005 than the weighted NVDRS-CPS sample. Confidence bands for the two curves generally overlap after 2005 (except in 2012), though the Army consistently has higher rates of suicide than the fully adjusted NVDRS-CPS sample. This suggests potential evidence of higher average rates in Army if one were to test across years rather than within years, as shown in the graphic.

In our analysis, we also identified five additional (unmatchable) factors—geography, parenthood, occupation, mental health, and firearm availability. These could be important when comparing Army with the general U.S. population depending on the type of question being addressed, with the needed caveats described earlier. However, we lacked the data needed to include these factors in our analyses.

We offer four recommendations, based on our assessment of data availability and our analysis of how different weighting factors for the general population affect the comparison between Army and NVDRS-CPS suicide rates.

1. **Given that comparisons will be made between the Army's suicide rate and that of the general population, those comparisons should adjust for age, gender, and year, and for the additional matchable factors of race/ethnicity, educational attainment, and marital status.**

As noted, accounting for factors such as race/ethnicity, educational attainment, and marital status notably shifted the estimated suicide rate for the NVDRS-CPS population in large part to changing the underlying Army population characteristics to which we are trying match the general population, affecting the conclusions one might draw from the Army-civilian comparison. As noted above, the decision to include marital status and education depends primarily on the underlying assumed theory about the relationship between Army service and these factors. If the theory is that the Army attracts individuals based on their education levels and marital status/aspirations, then both should be included in the matched comparisons between the Army and the general population. If education and/or marital status is being driven to change by DoD policies, it would be best not to include them directly in the matching because their inclusion might obscure the impact of serving in the Army on suicide risk.

2. **The Army should collaborate with the U.S. Census Bureau, the Centers for Disease Control and Prevention (CDC), and the U.S. Department of Labor to improve occupation/industry coding for general population deaths.**

A soldier's job-related duties and operational tempo ("unmatchable" factors) are other factors that may distinguish the Army from general populations. However, we were unable to draw parallels between general population and Army job categories due to limitations in how occupation is coded in the mortality data available on the general population. Given that occupation is a known risk factor for suicide in both populations, better quality data on the general

U.S. population would be useful to obtain. A collaboration between the Army, Census Bureau, CDC, and Department of Labor could increase the priority assigned to more accurate coding of occupation and industry in death records for the general population. Additionally, extensive work would be needed to decide how to determine which general population occupations best align with military occupations. Preliminary work in this area has been done by Wenger et al. (2017) and could be used as a basis for this work.

3. **The Army should collect voluntary data on soldiers who own personal firearms and should encourage the CDC or another federal agency to resume collecting voluntarily provided survey data on gun ownership and use in the general population.**

Soldiers may differ from their general population counterparts regarding ownership of or access to personally owned firearms, the suicide method used in the majority of Army suicides. Adjusting for this factor may also be important for making comparisons between the Army and general population. As noted, this factor falls into our third set of characteristics for which careful consideration is warranted before inclusion in the standardization process. If the goal of an analysis is to estimate the effect of Army service on suicide risk, including such controls as access to firearms should be avoided because this characteristic represents the specific mechanism by which Army service affects suicide risk (e.g., through availability of military firearms). However, there may be some research purposes for which one does want to control for firearm access, for example, to directly study how much of the effect of Army service on suicide risk is mediated through access to military firearms. Unfortunately, high-quality data in both the general population and the Army is fundamentally lacking. The lack of data on personally owned firearms among soldiers and the general populations impedes the Army's ability to adjust for or study a potentially important factor that may distinguish soldiers from members of the general population and that is correlated with suicide.

4. **Future research should examine the suicide risk among those with mental health diagnoses in the Army relative to similar individuals in the general U.S. population**.

The Army and general U.S. population may differ with respect to mental health conditions, which are among the strongest risk factors for suicide. For example, the 2014 Army Study to Assess Risk and Resilience in Servicemembers (STARRS) showed that lifetime prevalence estimates of a variety of mental disorders were significantly higher among new soldiers who were surveyed during their first few days after reporting for duty than similarly matched individuals from the U.S. population on age, gender, education, and race/ethnicity in 2011–2012. This highlights the underlying challenge in comparing the Army's suicide rate with the general U.S. population in that the kinds of people who joined the Army are generally at substantially higher risk of suicide related to history of mental disorders than similar individuals from the general U.S. population who could have, but did not, enlist. Even if these types of differences do not extrapolate to all years (e.g., it might also be that all those new soldiers with high burden of prior psychopathology never made it past their first year of service and did not contribute to the high Army suicide rate), there is a great need to be able to match the two populations on mental health to be able to better understand the role mental health plays in suicide rates and the comparison of rates between the Army and the general U.S. population. Data deriving from medical claims may be most easily linked to death data and have detailed information on mental health diagnoses and thus may be the most fruitful avenue for future research. The Army could replicate the methods used in this study to examine the rate of suicide among those with mental health diagnoses in the Army relative to individuals in the general population with the same diagnoses, adjusting for the sociodemographic characteristics described above. Such research will likely require partnership with an existing health system or data system like the National Inpatient System that not only reports mental health diagnoses within an insured population but also links mental health information to cause of death data. To do this, data on diagnoses will be needed not just for suicide cases but also for the entire Army and general U.S. populations at risk. Additionally, care will need to be taken to address any differences in general population and Army/military health systems regarding coding of psychiatric diagnoses.

Acknowledgments

We thank our sponsor, Les McFarland, for comments and guidance on this research. We also are grateful to Kristin Saboe and MAJ Calvin Hutto within the Army Resilience Directorate who provided invaluable help and thoughtful feedback for this study and to Nicole Chevalier who contributed to this project during her time as a RAND Army Fellow. Staff at the National Center for Injury Prevention and Control at the Centers for Disease Control and Prevention also provided generous assistance. We are thankful to Mary Vaiana for improving the quality and readability of the report. And, finally, we are thankful to Michael Schoenbaum at the National Institute of Mental Health for feedback on an earlier draft of this report and to our reviewers, Bonnie Ghosh-Dastidar at RAND and Ronald Kessler at Harvard Medical School, for their valuable guidance.

Abbreviations

AIC	Akaike Information Criterion
CDC	Centers for Disease Control and Prevention
CI	confidence interval
CMF	Career Management Field
CONUS	continental United States
CPS	Current Population Survey
DMDC	Defense Manpower Data Center
DoD	Department of Defense
DoDSER	Department of Defense Suicide Event Report
ES	effect size
GED	General Education Development
I&O	industry and occupation
MOS	military occupation specialty
NIOCCS	NIOSH Industry and Occupation Computerized Coding System
NIOSH	National Institute for Occupational Safety and Health
NVDRS	National Violent Death Reporting System
NVSS	National Vital Statistics System
OR	odds ratio
PS	propensity score
STARRS	Study to Assess Risk and Resilience in Servicemembers
TWANG	Toolkit for Weighting and Analysis of Nonequivalent Groups
WONDER	Wide-Ranging Online Data for Epidemiological Research

CHAPTER ONE

Introduction

Since 2012, suicide has claimed the lives of more than 40,000 people each year in the United States (approximately 13 people for every 100,000), making it one of the top ten leading causes of death (CDC, 2015). Like the United States more broadly, the U.S. Army has seen increases in suicides, losing over 100 active duty soldiers to suicide annually over the past three years (approximately 25 for every 100,000) (Pruitt et al., 2018).

The Army Resiliency Directorate oversees the Army's Ready and Resilient strategy, responsible for "strengthening individual and unit Personal Readiness and fostering a culture of trust" (U.S. Army, 2016b). One of several directorate functions is to provide commanders and leaders with awareness and analysis of suicidality in the Army. The Army Suicide Prevention Program tracks suicides in the Army, provides descriptive analysis of potential factors affecting suicidality, and makes recommendations to commanders and leaders for mitigating suicide risk. The Army bases its assessments of risk and protective factors on analysis and observations of soldiers and, to some extent, on general population studies of suicide.

Little is known about whether or how serving in the Army might alter suicide risk for those who choose to serve. To directly answer this question, one would need to know what the suicide risk for soldiers would have been had they not joined the Army, and that is unobservable. To approximate this, however, we can compare soldiers with individuals who did not join the Army but who are otherwise matched on observed risk factors for suicide. This comparison group would represent what soldiers' suicide risk might have been had they not joined the Army, conditional on the observed factors that went into the comparison. However, this approach cannot include the unobserved characteristics on which the two populations differ; therefore, it cannot prove that the kinds of people who join the Army are not different from those who do not join with respect to suicide risk factors or that Army experiences do not increase risk of suicide. It might well be that unmeasured covariates exist such that (for example) the kinds of people who join the Army are more "psychologically resilient" than individuals in the general population but that soldiers face much more severe stresses, resulting in the adjusted suicide rates (for the measured covariates we balanced on) being the same for soldiers as the general U.S. population, even though the suicide rate of soldiers would have been lower than for the general U.S. population in the absence of the unadjusted special stressors to which Army personnel are exposed.

Additionally, the implications of using different sets of factors in the process of matching the general population to the Army is not well understood. Broadly speaking, any comparison of Army suicide rates with the general population needs to be able to standardize on some core characteristics that (a) are associated with suicide risk, (b) differ between military and the general population, and (c) are outside the control of the Army. These factors include age, sex, and race/ethnicity. Without standardizing on these core demographic factors, any comparison between Army and the general population suicide rates will be affected substantially by underlying differences in the populations on these characteristics. So, matching on these is a needed first step. Conceptually, there are three different types of factors on which we might aim to match or standardize the general U.S. population to look like Army.

Set I includes demographic characteristics, such as age, gender/sex, race/ethnicity, and geography of origin. These factors represent stable characteristics that differ between the Army and the general U.S. population because those who choose (or are allowed) to enter the Army are not fully representative of the general population. As noted, these factors are critical to include in any suicide rate comparisons between the Army and the general U.S. population because they fulfill requirements (a), (b), and (c). Set I also includes a number of unmeasured factors on which we will not ever be able to match the two populations (e.g., psychological resilience), highlighting the inherent complexity in any calculations that aim to compare the suicide rates between the Army and the general U.S. population.

Set II includes characteristics that might be somewhat influenced by Army service, such as marital status and education. These factors are largely determined by individual service members' personal interests and aptitudes that generally predate their military service but may also reflect military policy or opportunities that occurred as part of military service.

Set III includes characteristics that are known to be directly affected by military policy or experiences, such as access to firearms, mental health, or occupation. Controlling for the characteristics in the third set needs to be carefully considered because matching on these types of characteristics would dramatically alter how one interprets any differences between the Army and the matched general population suicide rates. These factors satisfy (a) and (b) but not (c). For these types of factors, it may be of interest to understand the impact of making different policy choices, and these factors might well be subject to interventions in a way that age, sex, and race/ethnicity are almost certainly not. As such, factors in this third set should be studied in a different way. For example, if the goal of an analysis is to estimate the effect of Army service on suicide risk, including such controls as access to firearms and mental health should be avoided. This is due to these types of characteristics representing the specific mechanism by which Army service affects suicide risk (e.g., through availability of military firearms or mental health problems due to military trauma). An analysis that would attempt to match on these types of factors would obscure the fact that suicide risk is higher or lower for soldiers specifically because they joined the Army and gained access to these mechanisms. That said, there may be some research purposes for which one does want to control for such factors. For example, if one wants to examine suicide differences between the Army and the general population for a subgroup of individuals with specific mental health diagnoses or to understand how much of the effect of Army service on suicide risk is mediated through access to military firearms, then it may be important to match the populations on these factors. Notably, any differences in suicide rates between the Army and matched general population group in these types of analyses should be interpreted cautiously. The distinction between the causal effects of Army service versus factors associated with selection into Army service, which are multidimensional, dynamic, and never fully distinguishable, is technically irrelevant for a top-line comparison of suicide patterns. These effects—and what the Army might do about them—is what the Army Study to Assess Risk and Resilience in Servicemembers (STARRS) is designed to be able to address. There is no possible way that it could or even should be incorporated into regular Defense Suicide Prevention Office–type surveillance reports.

Characteristics in the second set also warrant careful consideration before inclusion in the standardization process. The decision to include factors, such as marital status and education, in Set II that might be somewhat influenced by Army service depends primarily on the underlying assumed theory about the relationship between Army service and these factors. For example, soldiers are more likely to be married than similarly aged individuals in the general population. One theory for this difference is that the Army attracts individuals who also have an interest in getting married (perhaps they are more religious or socially conservative than the general population, or they have a greater interest in having children). If this is the correct theory, marital status should be included in the matched comparisons between the Army and the general population in the same way as variables in Set I. Alternatively, if the higher marriage rate among soldiers reflects Department of Defense (DoD) policies designed to promote marriage under the belief that marriage makes for healthier soldiers, then matching on marital status will produces a comparison between Army and the general population that ignores the effect of Army policies designed to promote soldier well-being. Under this second type of theory, one should avoid matching on this factor when trying to estimate the effect of Army service on suicide risk in the same way that they should be careful with variables in the third set of factors. In short, there are some characteristics for which it is not completely clear whether they should or should not be controlled for in any comparison between soldiers and the general population. Those decisions will need to be informed by the researchers' and Army's theory about the relationship of such factors to Army service and the specific goals of the analysis.

While it is relatively easy for the Army to conduct analysis of suicidality within its own ranks, there have been no comprehensive efforts to study suicide in the general U.S. population that have risk and protective factors comparable with those of regular component soldiers. Army suicide research and programming would benefit from having a set of more accurate comparisons with the general population sector so that more robust comparisons of risk and protective factors can be made between the two populations.

The Army is often seen as a microcosm of the nation and as such is often compared with the nation and national statistics on measures that vary from the price of groceries to suicide rates. However, soldiers as a group differ demographically from the nation as a whole in ways that could be associated with suicide risk. There are, for instance, fewer women represented in the regular Army component proportional to the general population—overall, women have a suicide rate much lower than their male counterparts. This would mean that comparisons between suicide rates in the Army and in the general population would exaggerate the risk of suicide that soldiers might face due simply to the gender composition. A more relevant general population comparison group would be the subset of the general population with gender distributions like those found in the Army. To accurately compare soldiers' risk with that of the general population, we need to identify general population comparison groups that are well matched to many risk factors other than gender.

Study Rationale

Comparing the suicide rate between military and nonmilitary groups is a common epidemiologic investigation. For example, in RAND's original report on military suicide, *The War Within* (Ramchand et al., 2011), comparisons were made between the Army active duty suicide rate and that of the general population—including both crude comparisons and those that accounted for differences between the populations' demographic profiles. Policymakers, military leaders, researchers, and the media often ask those studying military suicide or working on suicide prevention programming how the rates compare.

Although it is rarely stated directly, there is a common rationale for asking this question. The Army is drawn from the general population of the United States. If rates differ between the Army and the larger population, then it signifies that one group is at greater risk of suicide. Adjustment via standardization—as we attempt to do in this study through comparison with a matched general population—aims to ensure that comparisons are not being explained by the matching factors themselves (e.g., age and gender differences in the population). Unfortunately, tackling such an analysis in a robust way is a complex task. The key problem in addressing the above questions is that the kinds of people who join the Army and the kinds of experiences to which Army life exposes a service member both change over time and are related to each other. The kinds of people who join the Army during times of economic downturn are different from the kinds of people who join the Army as a patriotic act in the wake of terrorist acts. Additionally, the experiences that are unique to Army service are different during times of war and times of peace. The research question would be much easier if the same selection mechanisms were at work during times of war and peace, as between-cohort differences in exposures to suicide experiences could be used to study the effects of time-varying experiences on time-varying within-service suicide rates. But that is not the case. An additional complexity exists in that we have strong selection out of service on the basis of risk factors for suicide. Close to 20 percent of new enlistees leave the Army in the first year of service, the majority of them for reasons that are related to risk of suicide, as the suicide rate among people with Army service is highest among people who leave the service prematurely in their first term of service.

The main contribution of this report is to assess available data sources for measuring observed suicide risk factors in the general U.S. population and the impact of standardizing the two populations using different sets of commonly available factors. We do not intend to estimate causal effects of serving in the Army. By understanding more fully how the comparisons shift depending on which set of factors are used to standardize the general population to look like the Army, we aim to show the Army the impact of different sets of adjustment factors. Suicide rates have increased in the Army, and the Army wants to understand how much might be explained by factors it can control and to gain insight into different databases that might be useful when making simple standardized comparisons between the two populations. It is inherent in all the provided estimates that unobserved factors are not matched on, and therefore the comparisons in suicide rates being shown all still have an important lingering bias that cannot ever be fully accounted for given real-world data limitations for both populations.

Organization of the Report

The purposes of this study were to (1) identify relevant general population databases that would be suitable for creating comparison proxies for the Army, (2) provide information on risk and protective factors for suicide in the general population that may be relevant to soldiers, and (3) analyze the degree to which suicide rates differ between the Army and the matched general U.S. population. This technical report surveys the data and methods available to improve how the suicide rates of the two populations are being compared.

Our discussion is organized as follows:

- In Chapter Two, we present the candidate factors that we identified as relevant for comparing Army and general population suicide rates. This chapter describes existing research on each factor and its relevance for the Army-general population comparison.
- In Chapters Three and Four, we describe the Army and general population data sources we used for our analyses, explain how we merged these data, and discuss what we learned with respect to the relationship between the factors identified in Chapter Two and suicide risk for each population independently.
- In Chapter Five, we discuss our approach to "matching" the Army and general population data and compare suicide rates over time given an expanded set of relevant factors on which to match the two populations.
- In Chapter Six we summarize our findings and present recommendations for the Army moving forward.

Chapters Three, Four, and Five are technical and suggested for readers interested in details about data sources and statistical methods to best match soldiers to the general population. We also provide a series of appendices that expand on points made in the main text.

CHAPTER TWO

Suicide Risk and Protective Factors

This chapter presents an overview of characteristics that differentiate Army and general populations and that are associated with suicide risk. Characteristics that meet these two criteria are candidate variables to use for comparing and studying Army and general population suicide rates. In the first section of this chapter, we review the scientific literature on these select suicide risk and protective factors, drawing from both the general population and Army literature. In the second section, we discuss the relevance of each factor to the Army-general population comparison and explain why we included it (or excluded it) in our matching analyses, which is discussed in Chapter Five.

Whereas the research on suicide risk among the general U.S. population derives from multiple sources, much of the most recent Army research reviewed here comes from the Army STARRS initiative. Army STARRS is a collaborative effort between the U.S. Army and the National Institute of Mental Health to investigate risk factors and protective factors for suicide, suicide-related behavior, and other mental/behavioral health issues in Army soldiers. It consisted of eight substudies that collectively involved compiling historical administrative data among more than 1.6 million soldiers on active duty; collecting questionnaire and neurocognitive data directly from more than 100,000 active duty soldiers; linking administrative and survey data; collecting blood samples from over 50,000 soldiers; and testing these samples for genetic and other biomarkers (DoD, 2018).

Overview of the Literature on Key Risk and Protective Factors for Suicide

Demographics

Gender

In the United States, men are approximately four times more likely to die by suicide than women. Women are much more likely to attempt suicide (Oquendo et al., 2001; CDC, 2015), but men represent nearly 80 percent of all successful suicides. The most viable explanation for the gender difference in suicide is that men tend to use more lethal (less reversible) means compared with women (CDC, 2015; Goldsmith et al., 2002). For instance, the most commonly used method of suicide among men is firearms, whereas the most common method for women is poisoning (CDC, 2015).

Suicides in the Army are also more common among men, and male soldiers are just over three times more likely to die by suicide than female soldiers. The Army STARRS study examined all regular Army members serving between 2004 and 2009; the suicide rate among female soldiers was 6.5 (per 100,000 person-years) compared with 20.4 for male soldiers (Schoenbaum et al., 2014).

Age

In the general population, the suicide rate generally increases from early adulthood through later life, at which point the suicide rate declines for women but increases for men (Schoenbaum et al., 2014). Figure 2.1 shows suicide rates from 2016 for men and women by age in the United States. The data shown reflect rates for individuals ages 18 to 75, the age range we expect to see in the Army. (We would expect few soldiers older than 58.)

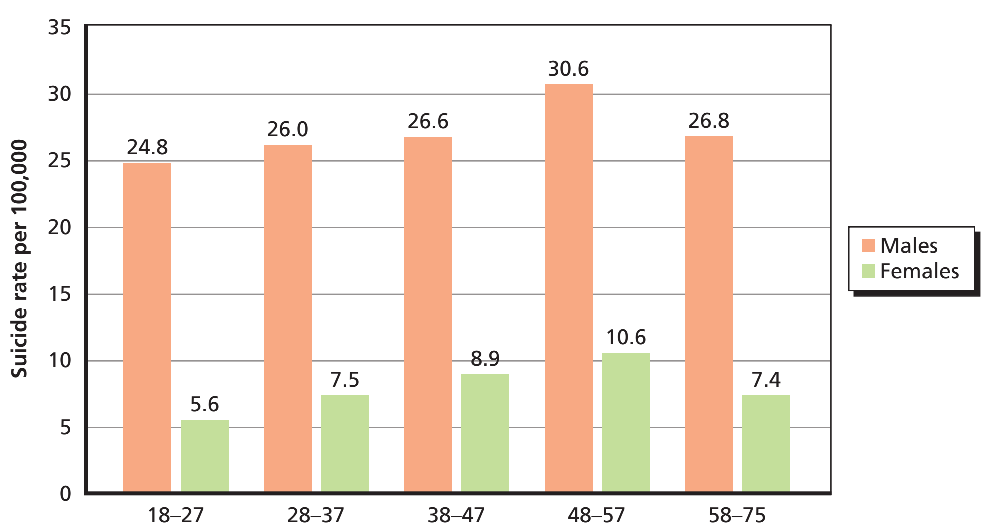
In the Army, suicide rates are highest among younger soldiers (those ages 17–20) with a rate of 23.4 per 100,000 people; rates generally decline over time to approximately 13 or 14 suicide deaths per 100,000 among soldiers over age 30 (Schoenbaum et al., 2014).

Race and Ethnicity

Overall, suicide rates in the United States tend to be highest among non-Hispanic whites and American Indian/Alaskan Natives (AI/AN) (Kochanek et al., 2016). In 2014, the suicide rate among non-Hispanic whites was 17.6 (per 100,000 people) and 10.8 among AI/ANs, compared with 5.6 for black non-Hispanics, 6.1 for Asian/Pacific Islanders, and 5.9 for Hispanics (Kochanek et al., 2016). However, there is some evidence that suicides are misclassified in ethnic minority groups; thus, reported suicide rates in these groups may be inaccurate (e.g., the lower rates among blacks and Hispanics may be underestimates) (Rockett et al., 2010).

Suicide rates in the Army follow similar patterns but tend to be higher among all race/ethnic groups relative to their general population counterparts, especially among Asian/Pacific Islander soldiers (Schoenbaum et al., 2014). In the Army STARRS study, the suicide rate among white and Native American soldiers was 20.2 and 35.9 per 100,000 person-years, respectively (Schoenbaum et al., 2014). The suicide rate among black soldiers was 12.8; among Hispanics, 16.0; and among Asian/Pacific Islanders, 21.5. However, such comparisons are not exact. The Army STARRS used race/ethnicity categories that were not mutually exclusive (i.e., a person

Figure 2.1
Suicide Rate for Males and Females by Age in the United States, 2016



NOTE: Data retrieved from CDC, "Fatal Injury Data," Data and Statistics (WISQARS) database, July 12, 2018b.

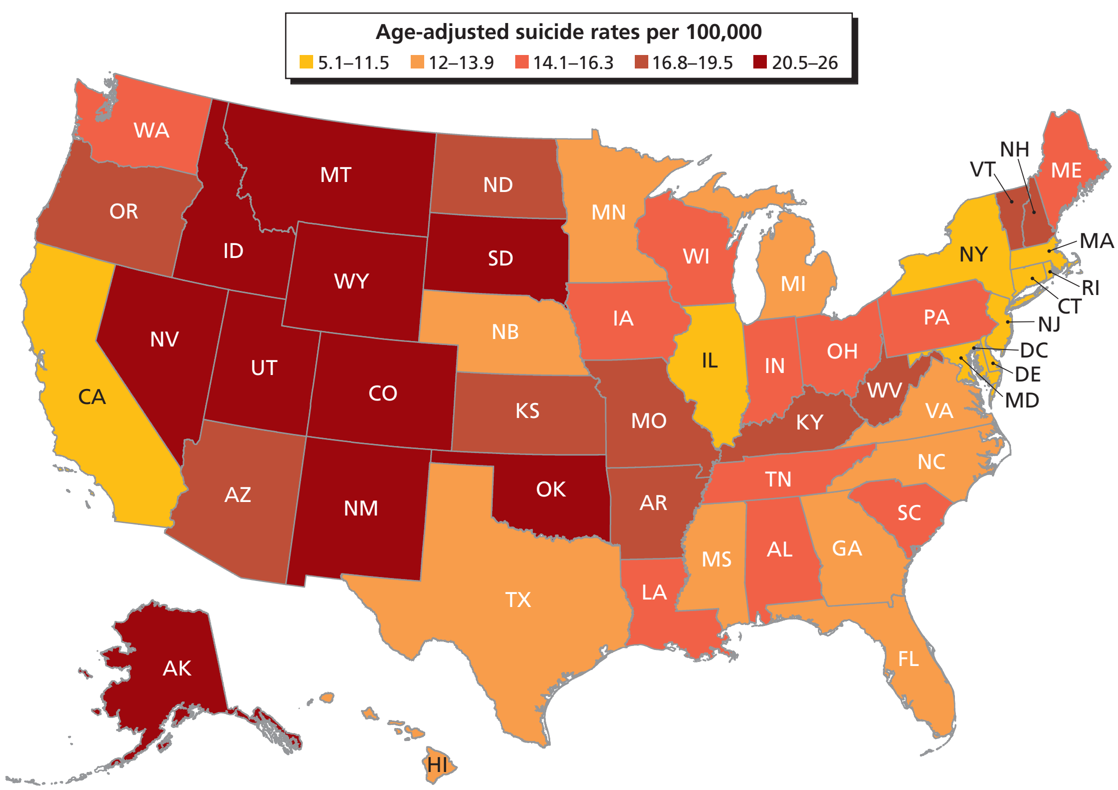
could belong to more than one race/ethnic group). Further, race/ethnicity category labels can differ between studies (e.g., Native American versus American Indian/Alaskan Native), and it is not possible to confirm that such categories would be interpreted the same way between participants in different studies.

Geography

Historically, the western mountain states and Alaska have had the highest suicide rates in the country (see Figure 2.2). This pattern has persisted for decades, even when controlling for race, age, and gender (Miller, 1980; CDC, 1997). Analyses of suicide rates have been conducted primarily at the state level; units such as county or census block help to measure variations in suicide risk that correspond to rural and urban areas or to other substate characteristics (Park and Peterson, 2014).

Research is mixed about whether regional variation in suicides reflects differences in population or differences in risk associated with location. It could be that local geographic or cultural factors elevate suicide risk for residents of the mountain west; in this case, the place itself is a risk factor. Alternatively, people prone to suicide may self-select to live in areas that have high rates of suicide, in which case population risk factors may explain elevated suicide rates in some states (Agerbo, Sterne, and Gunnell, 2007; Lester, 1995; Shrira and Christenfeld, 2010).

Figure 2.2
Suicide Rate by State, 2016



NOTE: Data retrieved from CDC (2018).

Regression analyses suggest that state variation can be partially explained by the following social and cultural factors:

- **education rates** (CDC, 1997; Price, Mrdjenovich, and Dake, 2009; Abel and Kruger, 2005)
- **access to lethal means** (Miller, Azrael, and Barber, 2012; Nock et al., 2008; Lubin et al., 2010)
- **economic factors** such as unemployment, income inequality, and poverty rates (Phillips, 2013; Milner et al., 2013; Cylus, Glymour, and Avendano, 2014)
- **measures of social cohesion** such as religion, marriage, and urban/rural differences (CDC, 1997; Smith and Kawachi, 2014; Agerbo, Sterne, and Gunnell, 2007; Phillips, 2013; Milner et al., 2013; Fontanella et al., 2015; Judd et al., 2006; Nock et al., 2008; Martin, 1984; Kposowa, 2013)
- **availability of health and mental health resources** (Lang, 2013; Tondo, Albert, and Baldessarini, 2006; Price, Mrdjenovich, and Dake, 2009)
- **alcohol/illegal drug use** (Phillips, 2013; Hourani et al., 2006; Johnson, Gruenewald, and Remer, 2009)
- **average elevation above sea level** (Haws et al., 2009; Cheng, 2010; Brenner et al., 2011).

We are not aware of research that has specifically examined geographic distribution of Army suicides in the United States. The Department of Defense Suicide Event Report (DoDSER), which is the official suicide surveillance system for the DoD, provides the following geographic detail: country at time of death, event setting (e.g., own residence, barracks, etc.), residence at time of event, and duty environment (e.g., garrison, leave, etc.).

Time and Seasonality

Between 1999 and 2014, the national suicide rate increased from 10.5 to 13.0 per 100,000 (Curtin, Warner, and Hedegaard, 2016). Most states and all geographic regions mirror the pattern of the national average (Stone et al., 2018). Rates across all levels of urbanization also mirrored this pattern between 2001 and 2015 (Ivey-Stephenson et al., 2017).

However, suicide rates over this period have not grown equally for all age, race, and gender groups. For example, suicide rates have shown a particularly strong increase for all men and women in the 45–64 age group and for white and American Indians of all age groups (Curtin, Warner, and Hedegaard, 2016; Ivey-Stephenson et al., 2017). The suicide rate of non-Hispanic whites ages 45–54, particularly those with low education, showed a particularly sharp escalation, from 21.8 deaths per 100,000 in 1999 to 38.8 per 100,000 in 2015 (Case and Deaton, 2015). Some of this increase may be tied to economic factors: suicide rates of 10-year age groupings between 25 and 64 rose with recessions and fell during economic expansions (Luo et al., 2011).

There are also consistent seasonal patterns in U.S. suicides in the general population. A recent review indicates that, with some exceptions, suicides peak in the spring, and a second, smaller peak appears in early summer (Christodoulou et al., 2012). We are not aware of any studies that examined seasonal variation in Army suicides. The Army's suicide rate has increased since 2000; we discuss this pattern in the chapters that follow.

Marital Status

A recent meta-analysis summarizes the literature on marital status and death by suicide (Kyung-Sook et al., 2017). Overall, suicide risk was higher for nonmarried versus married individuals (odds ratio [OR] = 1.92). However, there were some differences across gender and age groups. For instance, unmarried men of all ages were at increased risk, but women over 65 were not. Divorced individuals were at higher risk than any of the other marital status categories. The risk for unmarried individuals was higher for those under 65 than for the elderly.

The predominant theory linking marital status with suicide risk points to the increased social, economic, and emotional support and decreased social isolation that are associated with marriage. Further, the reduced risk might be due to selection bias with people less prone to commit suicide being more likely to marry.

In military samples, the ARMY STARRS study showed the same pattern: married soldiers and soldiers with dependents were at lower risk than unmarried soldiers without dependents (Schoenbaum et al., 2014).

Parenthood

Several studies suggest that parenthood confers a protective effect on women. The highest suicide rates among women occur in those who are childless; the rates decline as the number of children (Cantor and Slater, 1995; Hoyer and Lund, 1993) increases. Another study found that the protective effect of parenthood, when adjusted for demographic, socioeconomic, and psychiatric health factors, was only significant in women with three or more children. In this same study, the age of the child is also relevant—younger children have a higher protective effect (Qin and Mortensen, 2003). The presence of children may tie to the social integration theory of suicide prevention, deterring suicide because of the strong social bonds between mothers and children (Veevers, 1973). Though no comparable effect has been found for men (Conejero et al., 2016), the presence of dependents was shown to be protective as noted above in the section “Marital Status” (Schoenbaum et al., 2014).

We are not aware of research that has specifically examined parenthood as a risk or protective factor for suicide in the Army.

Occupation

Most research on general population occupation and suicide has been narrow, focused on a single occupation or a small set of occupations (Boxer, Burnett, and Swanson, 1995). However, several studies have examined variation in suicide risk across occupations. The studies have found farming, fishing, and forestry occupations at high risk for suicide death (Lavender et al., 2016; Stallones et al., 2013; Tiesman et al., 2015); a recent meta-analysis confirmed that agricultural, forestry, and fishery workers had about 1.5 times the risk of other occupations (Klingelschmidt et al., 2018). Another high-risk group identified in several studies is health care workers (Lavender et al., 2016; Stack, 2001; Stallones et al., 2013). A few studies have indicated elevated risk among firefighters and law enforcement (or protective services) (Tiesman et al., 2015; Stallones et al., 2013). Elementary school teachers may have lower than average risk (Stack, 2001).

However, studies that control for several demographic (chiefly age and gender) and educational factors (discussed below) show attenuated occupation effects. Overall, sociodemographic effects were seen as the largest contributor to the differential suicide risk across occupations (Bhatia, Rathi, and Kaur, 2014).

Other explanations of differential suicide risk by occupation focus on the predictors of suicide within an occupation. These studies tend to examine factors such as job-related exposures to risk factors, workplace characteristics, and demographic factors. For instance, posttraumatic stress disorder has been identified as a suicide risk factor among firefighters (Henderson et al., 2016), smaller police departments as a risk factor for police officers (Violanti et al., 2012), and pesticide poisoning as a risk factor among occupations commonly exposed at work (Stallones, 2006). Older male farmers are especially at risk among farmers (Boffa et al., 2017).

Early studies of suicide in the U.S. military focused on rank. The studies found higher risk among junior personnel, but they did not explore occupation (Helmkamp, 1995). Since then, several studies have found different risk levels among certain occupations, including elevated risk for infantry, gun crews, and seamanship specialists (with tactical operations officers at lower risk; Trofimovich et al., 2013); elevated risk among infantry and combat engineers (Kessler et al., 2015a); elevated risk among infantry or special operations within Army and Marines (Anglemyer et al., 2016); elevated risk for aircraft-related and other occupations in the Navy (Anglemyer et al., 2016); and elevated risk for aircraft-related, police, corrections, and firefighting occupations in the Air Force (Anglemyer et al., 2016). However, research findings are not consistent: One study found no link between occupation and suicide risk once the analysis controlled for age and gender (LeardMann et al., 2013).

Educational Attainment

The relationship between individual educational attainment and suicide risk is complex. The most recent analysis shows a nonlinear relationship for both men and women (although more pronounced for men): rates are lowest for those with a college degree and highest for those with a high school degree. However, those with less than a high school degree have a lower rate of suicide than those with only a high school degree (Phillips and Hempstead, 2017). This study also found that suicide rates increased between 2000 and 2014 for all educational attainment groups.

In the Army STARRS study, suicide exhibited a linear relationship with educational status. The highest rates of suicide occurred among those with less than a high school degree (20.8 per 100,000) and an alternative educational certificate (35.5 per 100,000), followed by those with a high school diploma or General Education Development (GED) (19.3 per 100,000), some college (11.8 per 100,000), and a college degree or more (9.8 per 100,000; Schoenbaum et al., 2014). The DoDSER for 2016 does not provide rates of suicide in the Army for educational attainment levels other than high school graduates (29.3 versus a total Army rate of 26.7 per 100,000) because fewer than 20 soldiers died within other strata (Pruitt et al., 2018).

Mental Health

One of the most robust predictors of suicide is a history of mental health problems. Here we discuss the relationship between mental health diagnoses and suicide death and between treatment for a mental health diagnosis and suicide death. We also describe how the availability of mental health treatment affects regional suicide rates.

Mental Health Diagnoses

Prospective studies of death rates among individuals with mental health diagnoses indicate that many disorders carry with them an increased risk of suicide. According to a meta-review (a systematic review of systematic reviews), the risk was tenfold higher than the general population for those with borderline personality disorder, depression, bipolar disorder, opioid use, and schizophrenia, and specifically for women with anorexia nervosa and alcohol use disorder (Chesney, Goodwin, and Fazel, 2014). There is variability across mental disorders and their association with suicide death. While many disorders are statistically associated with increased suicide risk, others question the clinical utility: even for disorders with elevated statistical risk, a low effect size coupled with the low absolute risk of suicide may still result in odds of suicide death close to zero (Bentley et al., 2016).

In addition, fewer than half of those with mental health conditions access mental health treatment (Wang et al., 2005). Thus, it is not clear whether those who died by suicide without a mental health diagnosis may actually have had a mental health condition that was not yet diagnosed. Some research has examined suicide risk postmortem and used interviews with family and friends to determine whether those who died may have had a mental health disorder (Kelly and Mann, 1996).

Army research on the suicide risk of mental health conditions has consistent findings. Risk for suicide was elevated among those with diagnosed mood disorders, anxiety disorders, posttraumatic stress disorder, personality/psychotic disorders, substance-related disorders, and adjustment disorder (Bachynski et al., 2012). Similarly, a psychological autopsy study conducted among 61 soldiers who died by suicide found that 70 percent had evidence of internalizing disorders (i.e., depression, mania, generalized anxiety disorder, posttraumatic stress disorder, and panic attacks) in the month before dying relative to 38 percent of matched controls (i.e., living soldiers who had demographic characteristics similar to those of soldiers who died by suicide). Externalizing disorders (e.g., substance use disorders, attention-deficit/hyperactivity disorder) were also common among suicide cases, though the difference in externalizing disorders between suicide cases (50.7 percent) and matched controls (36.3 percent) was not statistically significant (Nock et al., 2017). In total, 79 percent of Army suicide cases in that study had any mental health disorder, relative to 51 percent of matched controls.

Mental Health Treatment

Across studies, history of prior psychiatric treatment is the strongest predictor of suicide (Franklin et al., 2017). A 2002 review revealed that 20 percent of suicide cases had contact with a mental health care provider in the month before suicide (Luoma, Martin, and Pearson, 2002), and a 2014 study of health plan members found that 24 percent had a mental health diagnosis in the month before they died (Ahmedani et al., 2014).

Similar findings exist for the Army: 28 percent of soldiers who died by suicide between 2004 and 2009 had a mental health care visit in the four weeks before their death compared with 8 percent of controls (Ribeiro et al., 2017). These results are typically *not* interpreted as showing that mental health treatment has a causal effect on (i.e., increases risk of) suicide; however, they may suggest that mental health services could be improved. It should also be noted that most suicide cases do not access mental health treatment. However, almost all the soldiers in this study had accessed some type of general health care in at least the year prior to death (Luoma, Martin, and Pearson, 2002; Ahmedani et al., 2014), and half of soldiers who took their own lives had some kind of health care encounter in the month before their death (Ribeiro et al., 2017) These data suggest that suicide prevention in general health settings could also be improved. Like mental health diagnoses, the predictive accuracy of prior *psychiatric* treatment is still relatively poor (weighted area under the curve [AUC] = 0.67, standard error = 0.06) (Franklin et al., 2017).[1]

Availability of Mental Health Care

Regional availability of mental health care may also affect suicide rates. A recent review indicates that suicide rates are lower in areas with a greater density of mental health care providers (Morral, 2018b). There is also evidence that as states expanded access to mental health care via mental health parity laws, suicide rates declined (Lang, 2013).

Firearm Access

A recent RAND review of the scientific literature describes how access to firearms affects suicide risk (Morral, 2018b). The evidence can be summarized in four points.

First, the risk of suicide is greater among new owners of firearms compared with the general population. A study in Washington state showed that persons who died by suicide compared with similar, living members of the public were more likely to have lived in a home where somebody had a license to own a gun (Cummings et al., 1997). A second study in California revealed that the suicide mortality rate among persons who had acquired a new gun was higher than that of the general population, after adjusting for age and sex differences between the two groups (Wintemute et al., 1999).

Second, changes in gun ownership rates are correlated with changes in regional suicide rates. For example, Miller and his colleagues (2006) estimated that a 10-percent reduction in regional firearm availability is associated with a 2.6-percent reduction in total suicides. Another research team estimated that a 1-percent increase in individuals with firearms in a population would be associated with a 0.7-percent to 0.9-percent increase in suicides (Briggs and Tabarrok, 2014).

Third, people who die by suicide are more likely to live in homes where guns are present than matched comparison samples of people who die by other means (Kung, Pearson, and Wei, 2005), people who are alive in the community (e.g., Conwell et al., 2002), and people with mental illness who do not die by suicide (Brent et al., 1994).

Fourth, international evidence suggests that policy changes that reduced the availability of firearms were associated with reductions in suicides. In Israel, a policy requiring some members of the Israeli Defense Force to leave their firearms at base when returning home on the weekends led to a 40-percent reduction in suicides for young adult men (Lubin et al., 2010). A series of policy changes in Switzerland that reduced firearm availability (for example, by shrinking the army and increasing the cost of weapons to soldiers after separating from the service) also led to fewer suicides among young adult men (Reisch et al., 2013).

Although these patterns are convincing, necessary caveats preclude claims that the availability of firearms causes suicide (discussed in Morral, 2018b).

Army-Specific Factors

Many factors relevant to suicide risk in the Army have no immediate analog in the general population. We focused on two such factors our review: rank and deployment.

---

[1] Weighted AUC is a meta-analytic metric of diagnostic accuracy evaluation; estimates of 0.50 indicate chance prediction and 1.0 indicate perfect prediction. A weighted AUC of 0.67 is suggestive of "poor" prediction for a clinical diagnostic test per available guidelines (Ana-Maria Šimundić, "Measures of Diagnostic Accuracy: Basic Definitions," *Medical and Biological Sciences*, Vol. 22, No. 4, 2008, pp. 61–65).

As noted above, early studies on suicide in the U.S. military found higher risk among junior personnel (Helmkamp, 1995); more recent studies show a similar inverse relationship between suicide mortality and rank among enlisted personnel (Schoenbaum et al., 2014). In 2016, the suicide rate in the Army was 28.8 per 100,000 people for E1–E4s and 27.0 for E5–E9s. The rate was not calculated among cadets, officers, or warrant officers because fewer than 20 soldiers died within each group (Pruitt et al., 2018).

With respect to deployment, findings are mixed. In one representative sample of soldiers who served between 2001 and 2007, deployment had no effect on suicide mortality in analyses that adjusted for demographic and time-varying covariates, including demographic factors and rank and component affiliation (Reger et al., 2015). However, Army STARRS did find a relationship between enlisted Army soldiers who served between 2004 and 2009—specifically, that suicide risk was elevated among enlisted soldiers and after their first deployment (though not necessarily after any subsequent deployments) (Schoenbaum et al., 2014; Gilman et al., 2014).

Relevance of Each Factor for Army-General Population Comparisons

We had three criteria for selecting factors to include when comparing the Army and general population suicide rates: (1) the factors must be associated with suicide risk in either or both the general and Army populations, (2) the factors must be differentially distributed between the two populations, and (3) valid, relevant data are available for both populations.

Table 2.1 lists the factors we used in matching Army and general population data in our study (referred to throughout the discussion as "matchable factors") and indicates why we excluded certain factors (referred to throughout as "unmatchable factors").

Demographic Characteristics

The U.S. Army differs from the general population with respect to most demographic characteristics. The Army is approximately 83 percent male (U.S. Army, 2016a), whereas women comprise 51 percent of the general U.S. population (Howden and Meyer, 2011). Most soldiers are between the ages of 17 and 40; none is younger than 17, and there are few soldiers in older age groups (Schoenbaum et al., 2014; U.S. Army, 2016a). In contrast, according to 2016 estimates, approximately 23 percent of the U.S. population is under 18 years, and about 15 percent is over age 65 (U.S. Census Bureau, 2016).

With regard to the proportion of racial and ethnic minorities, the Army is similar to the U.S. population overall: racial and ethnic minorities comprise about 40 percent of each population (U.S. Army, 2016a; U.S. Census Bureau, 2017a). However, the Army has a larger proportion of blacks and a smaller proportion of Hispanics when compared with the general population (U.S. Army, 2016a; U.S. Census Bureau, 2017a).

Given the availability of demographic data (gender, age, and race/ethnicity) in both Army and general population data sources, the differences in the demographic composition of the Army and general population, and demographic differences in risk for suicide, we deemed it important to match on demographic characteristics.

Geography

The Army presents a complicated case for the influence of geography. Compared with the general population, soldiers are particularly mobile, moving frequently between states and

Table 2.1
Factors Considered for Use in Matching Army and General Populations

Factor	Available in Army Data	Available in General Population Data	Included in Matching	Reason for Exclusion
Gender	Yes	Yes	Yes	Not applicable (N/A)
Age	Yes	Yes	Yes	N/A
Race/ethnicity	Yes	Yes	Yes	N/A
Geography	Limited	Limited	No	Data sources limited for both; for Army, rarity suicides does not enable examination of regional variation with precision
Time	Yes	Yes	Yes	N/A
Marital status	Yes	Yes	Yes	N/A
Parenthood	Yes	No	No	General population data sources are limited
Occupation	Yes	Limited	No	General population data sources are limited
Educational attainment	Yes	Yes	Yes	N/A
Mental health	Limited	Limited	No	General population data sources are limited
Firearm access	Limited	Limited	No	Data sources limited for both; for Army, classifying access to firearms among soldiers presents challenges

overseas deployments. This raises the question of whether the relevant geographic risk factors concern the location that soldiers called home at the time of their death, where they were born, where they were stationed when they died, or where they spent the most time in their lifetime. The general population literature provides no information about which of these locations may embody the relevant geographic risk factor.

Moreover, bases do not always share the characteristics of their surrounding communities. There are differences in socioeconomic status, access to firearms, social isolation, culture, and demographics. In addition, some of the factors that are used to explain geographic variation in the general population's suicide rates are irrelevant for the military population. For instance, for the general population, geographic differences may be associated with different health care availability or unemployment rates.

In addition to these factors, as described in Chapters Three and Four, we have more comprehensive geographic data on Army suicides than we have on general population suicides for matching. The latter is restricted to a subset of states (though the subset will grow in future years). In our analysis, we include the entire Army sample and compare them to a subset of the general population from a subset of states; as such, it was not feasible to match on state or any smaller geographic region.

For these reasons, we do not include geography in our matching; however, as described in Appendix E, we conducted a sensitivity analysis restricting the Army sample to those in the continental United States (CONUS) (e.g., where they were stationed when they were at risk or died from suicide*) during a calendar year in order to assess whether restriction to Army soldiers being* *stationed in the United States has an impact on our suicide comparisons. We found that results were not sensitive to this restriction and, as such, utilize the entire Army sample in our analysis.*

Time and Seasonality

Different years show different risk factors and different suicide rates for both Army and general populations, although some of the factors that may influence the change in general population rates, such as unemployment and economic conditions, may be irrelevant to the Army.

Nonetheless, it is important that we match suicides between the Army and general populations according to year. We conduct our matched analysis within a calendar year.

Marital Status

The Army differs from the general population with respect to marital status: in 2016, 57 percent of soldiers were married whereas only 48 percent of the general population reported being married in the 2016 U.S. Census (50 percent if you include those married but separated) (U.S. Census Bureau, 2017a).

Given the availability of marital status in both Army and general population data sources, the differences in marriage rates between the groups, and the lower risk of suicide among married individuals, we deemed it important to study the impact of matching on marital status.

Parenthood

In 2016, 45.6 percent of active duty soldiers had children or dependents. There is no exact general population analog (i.e., dependents), although in 2017, 41 percent of U.S. families lived with their own children who were 18 or under, though this is likely an underestimate of the total number of adults with children or dependents (DoD, 2016; U.S. Census Bureau, 2017a).

The data source we use for general population suicides, the National Violent Death Reporting System (NVDRS), does not include data on parenthood status, and thus we do not match for it in our analysis.[2]

Occupation

Within the Army, data on industry and occupational categories are strong. Data are available for military occupation specialty (MOS) at the person level, across time. Though some MOS codes changed across time, it is possible to create a crosswalk, at least at a level that can be used in analysis.

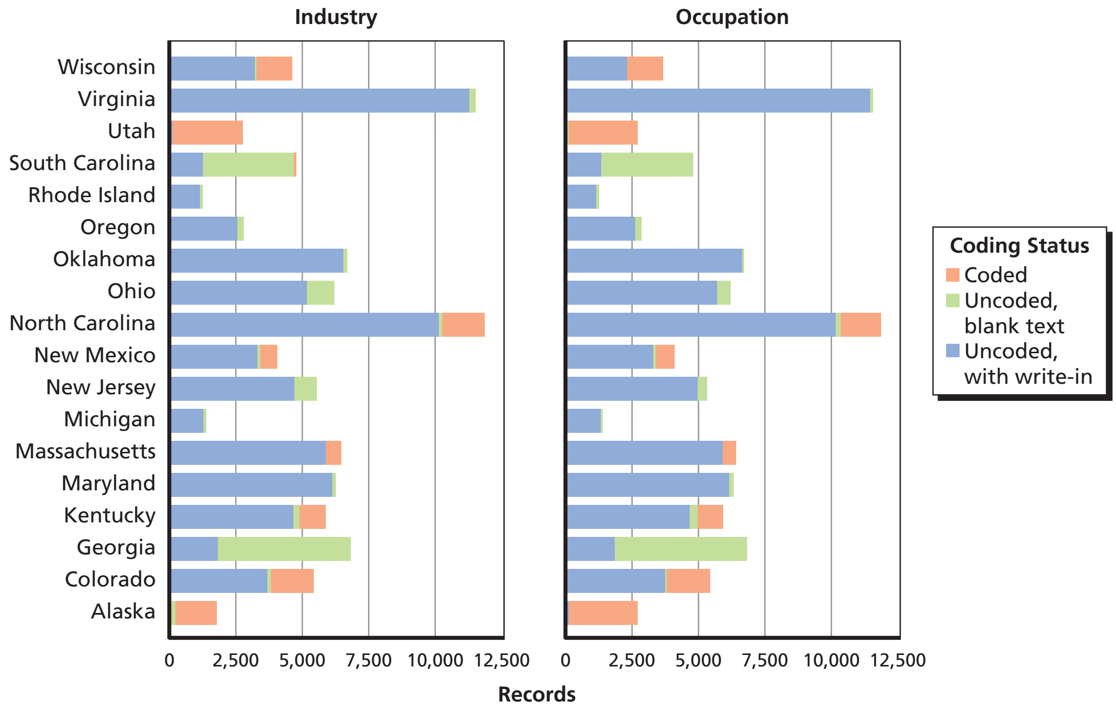
However, there are problems with general population data on occupation and industry. The NVDRS contains information on industry and occupation, but our analysis of the data suggests that these fields are often missing data (see Figure 2.3). When occupation and industry information is available, it is coded differently than it is in the Current Population Survey (CPS), the data set we use to construct our general population denominator to estimate risk by subgroup characteristics.

There is an additional challenge: even if better general population data existed, it would be difficult to compare military occupation codes with general population occupation codes. The jobs are very different, though there are some crosswalks between the two (Wenger et al., 2017). If better general population data on occupations and suicide were available, it might be

---

[2] Our reasons for selecting the NVDRS for civilian suicides are described in Chapter Four.

Figure 2.3
Provision of Census Industry and Occupation Codes by State, 2016 National Violent Death Reporting System (NVDRS)



possible to match on broad occupational classes, though the comparability of, for instance, "law enforcement" or "health care services" in the general population and Army settings is probably not great.

For these reasons, we do not include occupation or industry in our matching, though we consider occupation as an "unmatchable factor" in Chapter Three. We also describe in Appendix A our efforts to create better general population data on occupation and industry. In Chapter Six, we recommend that the Army partner with other relevant federal agencies to improve data on occupation/industry in general population mortality records.

Educational Attainment

The educational status of the Army differs from the general population, even when the comparison is restricted to those 18 and above. Specifically, 0.2 percent of the Army has not completed high school, relative to 11 percent of the general population, whereas 75 percent of the Army has completed high school or some college relative to 58 percent of the general population. A greater share of the general population has completed college (20 percent) or has an advanced degree (11 percent) relative to the Army (15.9 percent and 8.6 percent, respectively) (DoD, 2016; U.S. Census Bureau, 2017b).

Given the availability of educational attainment data in both Army and general population data sources, the differences in educational attainment between the groups, and the nonlinear relationship between educational attainment and suicide, we deemed it important to study the impact of matching on educational attainment.

Mental Health

Many of the same patterns between mental health and suicide in the general population are mirrored in the Army. However, matching soldiers with the general population on mental health status poses multiple challenges. First is the availability and validity of data on mental health disorders in the general population. General population data from the NVDRS indicates whether the decedent was known to have had a mental health condition, as ascertained from death certificates; coroner/medical examiner reports (including toxicology); and law enforcement reports, often derived from interviews with family and friends. We are not aware of any study examining the validity of these reports, and because mental health conditions are often untreated or unknown to family and friends, we suspect that they underrepresent the presence of mental disorders, including those that may be undiagnosed.

Second is the availability of Army data on mental health status. It is feasible to link Army suicide decedents to TRICARE health records to examine the proportion with a disorder and, more importantly, the risk of suicide among all such persons with a mental health diagnosis. While possible, accessing the data requires significant administrative work. Given our concerns about the reliability of general population data, we did not pursue the Army data.

For these reasons, we do not include mental health conditions in our matching. However, we make a recommendation in Chapter Six to specifically compare suicide risk among Army soldiers with mental health diagnoses to individuals in the general population with diagnoses in other health systems, adjusting for many of the differences we discuss in this chapter.

Firearm Availability

When considering whether it is possible to match the Army with the general population on firearm availability, we were presented with three challenges. First, there is variability in firearm access in the Army. We examined whether such access varies by occupation by reviewing descriptions of each Army MOS (taken from the "*MOS Smartbook*"); however, we realized that the descriptions were not valid for our purposes. For example, an MOS like 91F (Small Arms Repair) is described as involving firearms because soldiers in this occupation work on firearms, but special operations MOSs (which carry and use firearms) are not described in this way.

Second, even if service members had occupations that did not typically require access to a firearm, many would nevertheless own personal firearms. We are not aware of any Army data source that tracks personal firearms owned by soldiers (who may, in fact, be prohibited from doing so), though there may be registries of soldiers living in military housing with firearms. Nevertheless, the 2016 DoDSER data indicate that 67.9 percent of suicides in the Army involve a firearm; of these, 92.6 percent used a personally owned versus military-issued weapon (Pruitt et al., 2018). Therefore, we do not believe we can identify firearms availability for soldiers in a way that could be used to match on this factor with general populations.

The third challenge is that there is limited information on firearm ownership among either suicide decedents or in the general population. We are not aware of any data set that includes an indicator of access to firearms among suicide decedents. The most up-to-date information among the American population that owns guns comes from the General Social Survey (from NORC at the University of Chicago), which provides information about household ownership for the country and for each of nine census regions.

For these reasons, we do not include firearm access in our matching.

Conclusion

We identified six factors that (1) are related to suicide in the general population and/or the Army, (2) differ in frequency between the two populations; and (3) have data available for comparing the Army with the general U.S. population. These factors, which we refer to as "matchable factors," are gender, age, race/ethnicity, time, marital status, and educational attainment. Traditional analyses comparing Army with general population suicide rates have considered the first three; race/ethnicity has been considered on occasion. Marital status and educational attainment have not, to our knowledge, been used to compare the two groups. In Chapter Five, we will explore the impact of including these additional factors in matched comparisons between the general population and the Army.

We also identified five additional factors ("unmatchable" factors) that could be important for comparing the two groups; however, lack of data limits our ability to include them in the analyses. These factors are geography, parenthood, occupation, mental health, and firearm availability. We explore the role of the first three in explaining suicide variation in the Army since we had some information on geography, parenthood, and occupation in the Army data. In contrast, these factors are not available in the general population data and as such will be factors that we cannot include in our matching. We consider the potential impact such factors might have on the comparisons between Army and general population suicide rates in the Army analyses presented in Chapter Three as well as in the matching analyses presented in Chapter Five.

CHAPTER THREE

Army Risk Factors

Our study's ultimate goal to help the Army determine if Army service itself should be considered a risk factor for suicide or if differences in risk between soldiers and the general U.S. population reflect differences in other, more general risk factors. In pursing this goal, we compare the general population and Army suicide rates after adjusting for an expanded set of risk factors in the two populations. In this chapter, we describe the Army data sources used in our study and discuss how both the matchable and unmatchable factors reviewed in Chapter Two explain variation exclusively within Army suicides.

The objectives for our analyses are (1) to understand (and confirm) the extent to which the matchable factors explain variation in suicide risk in the Army and (2) to understand how unmatchable Army-relevant factors—such as geography (measured by unit location), military occupation, parenthood, and other characteristics unique to serving in the Army (deployment history and rank)—explain variation in Army suicide risk above and beyond matchable factors. While the Army-relevant unmatchable factors differ from the set of unmatchable factors identified as part of our literature review in Chapter Two (geography, parenthood, occupation, mental health, and firearm availability), the analyses of Army-relevant unmatchable factors are important for helping to inform the potential role omitted variables might have on the reported suicide rate comparisons reported in Chapter Five.

Our analysis has two steps. First, we develop a model to explain the association of suicide risk with matchable factors (i.e., gender, age, race/ethnicity, time, marital status, and educational attainment). This baseline model allows us to generate predicted suicide odds as a function of only matchable factors. Second, we develop additional models to help understand the extent to which suicide risk among soldiers varies by unmatchable factors that are available to us in the Army data (i.e., rank, deployment history, unit location), even after controlling for the baseline predicted suicide risk from the six matchable factors. These additional models help us determine the potential role the unmatchable factors might play in biasing conclusions when comparing Army with general population suicide rates.

The unmatchable factors are basically unobserved factors that we cannot control for when comparing Army and general population suicide rates. It is important to consider the potential impact they may have on our comparisons. Thus, for example, if rank is not associated with suicide after we have adjusted for risk due to the matchable factors (e.g., the association of suicide with rank is driven entirely by how rank relates to age and education), then the fact that we cannot match on rank in the comparison between Army and general populations is likely not a serious limitation. On the other hand, if suicide risk is uniquely associated with rank in the Army, this suggests that comparisons between Army and general population suicide rates might vary across Army subgroups with differences in suicides being higher versus lower for different ranked groups in the Army.

Data

RAND Arroyo Center received Army data for 2003 to 2015 from two sources. Suicide data were provided by the sponsor, who obtained the data from the Office of the Armed Forces Medical Examiner. All other data on soldiers were acquired from the Defense Manpower Data Center (DMDC). The DMDC is a comprehensive, central archive of personnel data housed at the United States Department of Defense, comprising interrelated databases that include demographic information, service history, geography (i.e., unit location), and occupational history (DMDC, 2006). This project obtained approval to use the Defense Enrollment Eligibility System (DEERS), the Activation and Deployment File (ADF), the Work Experience File (WEX), active duty pay files, and reserve pay files (though this last data source was not used). We received RAND Institutional Review Board approval as well as second-level approval from the Army.

Table 3.1 describes these data sources and the variables they contain that could be used to examine variation in suicide risk among soldiers.

Data Merging and Cleaning

We merged the DEERS, WEX, pay, and ADF files at the person-month level (i.e., each individual had data from each source at each month that he or she was in the Army) using scrambled social security numbers—unique longitudinally consistent identifiers for individual soldiers generated by DMDC.

Data Lagging

Army data are populated within each of these data sources at different times of the month (which we term "data snapshots"), resulting in some incompatibilities between the different data sources. For example, a soldier who dies by suicide in early January may not be represented in an administrative data source that takes its "data snapshot" at the end of the month, because that person is no longer alive. Due to the timing of administrative data snapshots for certain files occurring at the end of a month, we used a one-month lag to ensure that all

Table 3.1
Sources of Army Data, 2003–2015

DMDC File	Frequency	Variables
DEERS	Monthly	Age, sex, race/ethnicity, education level, marital status, number of dependent children, unit location
ADF	Monthly	Deployment history, location codes
WEX	Monthly	Primary service occupation, pay grade, deployment history, component, cumulative months in regular component; derived variables included cumulative months of deployment and number of deployments
Active Duty Pay	Monthly	Monthly total pay, basic pay
Reserve Pay	Monthly	Not used in this study

suicide cases were correctly merged into the *year* in which their death occurred. To perform this lagging procedure, we coded the December administrative data from the prior year (representing the soldier's characteristics at the end of December) as January of the target year, and January through November are incremented forward one month to be coded as February through December of the target year. Thus, our data represent the members' characteristics at the beginning of the coded month, rather than the more standard practice of representing the end of the month.

Missing Data

There were some, though minimal, missing data for gender, race, and age (range = 0.4 percent to 3.5 percent). Longitudinal, within-year filling was used as a first step to address these missing data across person-month records; in other words, there may sometimes be a lag between when a soldier is first represented in the data and when all of the soldier's demographic information becomes available. For example, if a person is missing information on gender in February but gender is not missing in March, we "fill in" the February value with the March value.

Following the longitudinal within-year filling, our data had very few missing values for demographic variables (range = 0.4 percent to 1.2 percent). As such, we handled missing data by coding those individuals with missing values of a covariate as part of the same group as those in the most frequently observed group from our data. Thus, for education, individuals with missing data were include as part of the "less than bachelor's degree" group; for race, the non-Hispanic white group; for age, the median age of 26; for marriage, the currently married group; and for gender, as part of the male group.

Creating a Person-Year File

As described below, to reduce computation time, we conducted most of our analyses on a simplified, person-year data set. To create the person-year data set, we used the value of covariates for each individual from the earliest observed record for each person within that year, conditional on the individual having served six or more months. Because of this, the Army sample was restricted to include only regular component personnel who had been in the Army for at least six months to ensure that they had at least one-half year to contribute to their first person-year.

Career Management Field

Part of our planned analysis was to examine variation in suicide rates by occupation in the Army. To do this, we use the Army Career Management Field (CMF), which groups soldiers into broader occupation categories—for example, infantry, engineer, or aviation. CMF grouping is typically determined by MOS codes. However, over the course of our time frame (2003–2015) there have been some major shifts and MOS code changes. For example, some Army intelligence MOS codes used to begin with 98, but those same positions recently changed to MOS codes beginning with 35. To accurately compare CMFs across time, we had to ensure that all the MOS codes were standardized to a certain year.

We standardized all MOS codes to 2015 for this study by using a combination of open source information and DMDC data. We identified any known recodes and shifts from Army sources or documents from 2003 through 2015. Concurrently, using the person-month data from DMDC, we identified large shifts in soldiers' MOS codes from one month to the next (not counting recruit MOSs). We looked at each soldier's MOS in one month and the following month. For each MOS, we identified the number of soldiers who remain in the same MOS, who switched into a new MOS, and who left the sample. We identified a potential recode/shift as 70 percent or more of the soldiers in an MOS moving into a new MOS the following month. Using the documented shifts and recodes, we validated the changes identified using only the DMDC data.

Our methods successfully identified the documented changes we had found and gave us confidence to use this method to identify the changes in years where we found no documented sources. Using a record of each year by month change we identified, we created a 2015 standardized MOS variable from which to group soldiers into CMFs and run analysis on occupational variation in suicide rates.

Erroneous Data and Final Missing Data

To identify erroneous or impossible item values, we first calculated basic descriptive statistics. For a few cases with age outside the expected range (i.e., less than 17 years or greater than 80 years), we imputed the correct age based on that individual's data elsewhere in the merged data set (n = 9) or assigned the age value to missing (n = 1). We also dropped any person-years that were present in the data following that person's suicide death (n = 15). Finally, we dropped 6,015 (0.08 percent) person-year observations outside the age range of 18 to 75.

Our final sample thus contained 7,249,908 person-year records among 1,355,535 distinct individuals.

Analysis

Our approach to modeling suicide risk in the Army population had two parts:

1. a baseline analysis that estimated the association between general population matchable factors and suicide in the Army
2. an extended analysis that estimated the association between unmatchable Army-relevant factors and suicide, after controlling for predicted suicide risk based on the baseline analysis.

There are several ways to model the association between these factors (both matchable and unmatchable) and suicide, which we discuss further in Appendix B. Based on our review of the methods described in the literature, we determined that the logistic link function would offer reliable estimates and allow greater comparability with prior Army STARRS studies (Kessler et al., 2015b; Gilman et al., 2014; Schoenbaum et al., 2014). Further, given the computing time it takes to estimate person-month regression models using data spanning 13 years (2003–2015), we also compared findings from person-month versus person-year covariate and suicide risk regressions. After finding very similar results, we opted to use person-year covariates as the primary predictors of suicide risk in the Army population.

The final regression modeling involved two steps.

Step 1: Estimate Baseline Model, Including Only General Population Matchable Factors

For the first step (the baseline model), we studied the associations of matchable factors with suicide deaths in the Army data. The matchable factors—those for which we can match the

Army population to the general population in studies of suicide outcomes—are gender, age, race/ethnicity, education level, marital status, and time. Our baseline logistic regression model can be expressed as follows:

$$\Pr\left(Y_{it}=1\middle|X_{it}\right)=\frac{\exp\left(\alpha_1+X_{it}\boldsymbol{\beta}\right)}{1+\exp\left(\alpha_1+X_{it}\boldsymbol{\beta}\right)}$$

where Y_{it} is a binary variable equal to one if individual i died of suicide by the end of year t (zero otherwise), $\alpha_1$is a constant, $\boldsymbol{X}_{it}$ is a vector of individual-level covariates for individual i at the beginning of year t, and $\boldsymbol{\beta}$ is a vector of estimated regression coefficients for the matchable factors in the regression model (age, gender, race/ethnicity, education, marital status, year dummies) corresponding to each covariate.

We explored how to code covariates within this model to ensure that we not only modeled the important associations but also limited the problems that can occur when using overly complex specifications to model rare outcomes. To begin, we specified the baseline model using very detailed covariate categories. Then, through standard model building exercises, we reviewed the coefficients to evaluate potential ways to collapse covariate categories with similar regression coefficients to estimate a more parsimonious model that would align with the covariate categories in the general population data and improve model fit, as assessed by Akaike Information Criterion (AIC) statistics.[1] We then iteratively added pairwise interactions between baseline covariates to our baseline specification to see whether any interactions improved model fit. Only one pairwise interaction (education level with age group) improved model fit and was therefore included in our final baseline model.

The final baseline model included the following covariate categories:

- **gender**: male; female
- **age**: age groups 18–20; 21–24; 25–29; 30–38; 39–44; and 39+
- **race/ethnicity**: non-Hispanic white; non-Hispanic African American; Hispanic; non-Hispanic Asian/Pacific Islander; other/unknown/not available
- **education level**: less than a bachelor of arts (BA); BA; more than college degree
- **marital status**: never married; formerly married; currently married
- **age by education interactions**: BA or higher education interacted with the following age groups: age groups 25–29, 30–38, and 39+
- **year**: dummy variables for each year 2003 to 2015

Using this model, we computed the predicted log-odds of suicide for each soldier in each year.

1 The AIC, an estimator that discriminates among models using Kullback-Leibler information criterion, provides a relative measure of parsimony to inform model selection. AIC rewards goodness of fit (likelihood function) and penalizes overfitting (number of parameters), with lower AIC values indicating a preferred model. Following guidance on thresholds for determining medium or large improvement (Kenneth P. Burnham and David R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach*, 2nd ed., New York: Springer, 2002; Kenneth P. Burnham and David R. Anderson, "Multimodel Inference: Understanding AIC and BIC in Model Selection," *Sociological Methods & Research*, Vol. 33, No. 2, 2004, pp. 261–304), we use changes in the AIC to select the covariate categories and interaction terms to retain in our baseline model.

Step 2: Estimate the Association Between Suicide Risk and Unmatchable Army-Relevant Factors

When estimating the second-stage model, we included for each soldier in each year their predicted log-odds of suicide from the baseline model as an offset (i.e., constraining its coefficient value to 1). This ensures that all the effects from the covariates in the baseline model are removed from the data before estimating the second-stage model. Thus, the second-stage model is investigating whether variation in any particular unmatchable factor explains risk in the Army even after accounting for the matchable factors. Here unmatchable factors included those Army-specific characteristics for which a general population counterpart will not have a measurement (e.g., rank) or those for which there may be a general population counterpart, but general population data are unavailable to make comparisons (e.g., located in a non-U.S. location).

Specifically, we estimated

$$\Pr\left(Y_{it}=1 \middle| Z_{it}\hat{Y}_{it}\right)=\frac{\exp\left(\alpha_2+Z_{it}\gamma+\hat{Y}_{it}\right)}{1+\exp\left(\alpha_2+Z_{it}\gamma+\hat{Y}_{it}\right)}$$

where α_2 is a constant, Z_{it} is a single unmatchable covariate (e.g., rank, non-U.S. location), γ is an estimated regression coefficient, and $\hat{Y}_{it}$ is the predicted log-odds of suicide estimated from the baseline model. In essence, this model allows us to evaluate the extent to which model fit is improved by including the unmatchable factor, and this in turn tells us how consequential for establishing a meaningful general population comparison group it may be that we cannot identify a general population counterpart who are comparable on the unmatchable factor.

We approach these analyses by examining unmatchable factors one by one to understand the potential impact each factor might have on explaining suicide risk in the Army population. In cases where there is no evidence of an association between an unmatchable factor and suicide risk above and beyond that which is explained by $\hat{Y}_{it}$ (baseline risk controlling for matchable factors), then we can view the lack of information in the general population as inconsequential for identifying differences in general population and Army risk. In contrast, factors that are found to be significantly associated with risk in the Army above and beyond $\hat{Y}_{it}$ indicate that the unmatchable factor is important for explaining risk in the Army, and that accounting for differences in the factor (i.e., stratification or subsetting) may reveal additional differences in comparisons with the general population suicide rate. More specifically, accounting for differences in the factor through stratification or sample subsetting (i.e., by effectively weighting one subset of the Army to zero) may affect the magnitude of the difference in suicide rates between matched Army and general population comparisons.

We used the following categories for unmatchable covariates:

- **geography**: unit location not in the United States, unit location in the United States but in non-NVDRS state, unit location in U.S. NVDRS state
- **rank**: E0–E2, E3, E4, E5, E6, E7–E9, O0–O2, O3, O4+, Warrant Officer
- **deployment history**: never deployed, deployed once, deployed two or more times
- **number of children**: no children, one child, two children, three or more children
- **CMF**: thirty-six occupational categories based on CMF (see Figure 3.1).

With the exception of CMF, the effects for the unmatchable covariates were modeled as fixed effects (i.e., estimated directly without assuming variability and interpreted as standard regression coefficients). Given our interest in a more detailed understanding of the role of occupation in explaining variation in suicide, combined with large differences in the size of different Army CMFs, we modeled CMF categories as random effects to reduce sample-to-sample variability through partial pooling of information across groups (Clark and Linzer, 2015; Lavielle, 2015). The random intercepts for CMF are not estimated directly but are characterized by their estimated variance.

Results

Population Characteristics

Table 3.2 shows trends in demographic characteristics of the Army sample from 2003 to 2015. There is a nonlinear relationship between time and population size, with the number of soldiers generally increasing from 2003 to 2011 (with exceptions in 2005 and 2010) and then decreasing steadily from 2011 to 2015. Starting in 2004, suicide rates began to steadily increase; rates appeared to stabilize between 22 to 23 per 100,000 persons starting in 2010, except for a sharp uptick in 2012 and then slight drop in 2013.

In terms of demographics, the sample does not change much over time. The sample is consistently predominantly male, with females constituting between 13.2 percent and 15.4 percent of soldiers. The sample is also predominantly non-Hispanic white, typically representing between 50 percent to 60 percent of the sample for a given year. Non-Hispanic African American individuals represent between 20 percent to 25 percent of the sample, and Hispanic individuals between 10 percent to 13 percent. The average age of soldiers ranges between 28 and 29, with the median age slightly lower at 26.

In all years, high school degree or equivalent was the most common education level, although the share of individuals with this level of education has been declining over the past five years. Overall, the sample seems more broadly distributed across education levels over time. In 2003, 4.8 percent of soldiers had less than a high school education and 17.8 percent had a BA or higher; by 2015, those shares had risen to 10.4 percent and 23.0 percent, respectively. Finally, the majority of the sample is "currently married," with rates ranging from 55 percent to 60 percent over the study period.

Baseline Model for Matchable Factors

Table 3.3 presents findings from our final baseline multivariate model, which includes the matchable factors. Most of the results confirm associations already seen in the Army (Schoenbaum et al., 2014) and described in Chapter Two. As expected, female soldiers tend to have lower rates of suicide (OR = 0.36; 95 percent confidence interval [CI] = 0.29, 0.46). Relative to non-Hispanic white soldiers, non-Hispanic African American and Hispanic soldiers have significantly lower odds of suicide (OR = 0.76 and OR = 0.72, respectively), while Asian/Pacific Islander and other racial/ethnic groups exhibit significantly higher odds (OR = 1.21 and OR = 1.45, respectively). Odds of suicide are higher for younger age groups, with a significant reduction in suicide risk occurring for individuals aged 30 to 38 (OR = 0.72; 95 percent CI = 0.58, 0.88) and 39 or older (OR = 0.59; 95 percent CI = 0.44, 0.78). Individuals with a four-year college degree or more have significantly lower odds of suicide than those with less education

Table 3.2
Descriptive Statistics for Army Sample Characteristics, 2003–2015

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
N	521,367	528,034	522,850	530,092	557,992	574,387	593,801	592,916	599,487	587,813	573,013	546,368	521,788
Suicide rate per 100,000	11.9	9.3	11.3	15.5	16.1	19.2	23.9	22.9	22.5	27.6	20.9	22.5	22.0
Gender													
Female	15.4%	15.1%	14.5%	14.2%	13.8%	13.4%	13.2%	13.5%	13.6%	13.5%	13.5%	13.8%	14.1%
Race													
Non-Hispanic white	59.9%	60.5%	61.4%	62.1%	60.8%	53.8%	51.5%	52.1%	50.6%	61.7%	61.1%	59.8%	58.7%
Black	24.4%	23.3%	22.2%	21.1%	20.3%	19.8%	19.5%	19.7%	19.9%	20.1%	20.4%	20.9%	21.4%
Hispanic	9.7%	10.1%	10.4%	10.5%	10.5%	10.6%	10.7%	11.1%	11.3%	11.6%	11.9%	12.4%	13.0%
Asian/Pacific Islander[a]	2.6%	2.9%	3.2%	3.6%	5.9%	13.1%	15.7%	14.6%	15.5%	4.3%	5.4%	5.7%	5.8%
Other	3.5%	3.1%	2.9%	2.7%	2.5%	2.6%	2.7%	2.6%	2.8%	2.4%	1.2%	1.2%	1.2%
Mean age	27.9	27.9	28.0	28.1	28.1	28.2	28.4	28.5	28.7	28.9	29.0	29.0	29.0
Education level													
<High school (HS)	4.8%	4.7%	5.0%	5.7%	6.0%	6.5%	6.2%	6.6%	7.5%	7.0%	11.1%	10.9%	10.4%
HS/GED	72.5%	72.5%	71.8%	71.3%	72.2%	72.0%	72.6%	71.6%	70.7%	70.3%	64.8%	63.2%	62.1%
Some college	4.9%	5.0%	4.8%	4.7%	4.5%	4.3%	4.0%	4.0%	4.0%	4.1%	4.2%	4.4%	4.6%
BA	12.0%	12.3%	12.6%	12.4%	11.9%	11.7%	11.7%	11.9%	12.0%	12.5%	13.4%	14.5%	15.3%
>BA	5.8%	5.7%	5.8%	5.9%	5.5%	5.5%	5.5%	5.7%	5.9%	6.2%	6.6%	7.1%	7.7%
Married													
Never	40.4%	39.2%	38.8%	38.0%	36.6%	36.0%	35.0%	34.4%	33.9%	32.6%	33.1%	33.7%	34.2%
Formerly	4.5%	4.8%	5.0%	5.3%	5.4%	5.6%	5.8%	6.0%	6.0%	6.2%	6.1%	5.8%	5.6%
Currently	55.1%	56.0%	56.2%	56.7%	58.1%	58.4%	59.2%	59.6%	60.1%	61.2%	60.8%	60.5%	60.2%

[a] We observed a seeming anomaly in the data whereby an unusual proportion of Army soldiers were coded as Hawaiian/Pacific Islander in 2007–2011. All analyses retain race/ethnicity classification based on the DEERS variable; however, this seeming difference in coding or reporting should be kept in mind.

Table 3.3
Baseline Multivariate Logistic Model of Association Between Suicide and Matchable Characteristics for Army Sample

	Adjusted OR	95% Confidence Interval (CI)
Gender (Ref: male)		
Female	0.36***	[0.29, 0.46]
Race/ethnicity (Ref: non-Hispanic white)		
Black	0.76*	[0.66, 0.89]
Hispanic	0.72**	[0.60, 0.88]
Asian/Pacific Islander	1.21*	[1.01, 1.44]
Other/unknown	1.45*	[1.06, 1.97]
Age group (Ref: age 18–20)		
Age 21–24	0.89	[0.75, 1.06]
Age 25–29	0.97	[0.80, 1.17]
Age 30–38	0.72**	[0.58, 0.88]
Age 39+	0.59***	[0.44, 0.78]
Education level (Ref: less than BA)		
Four-year college degree (BA)	0.20***	[0.09, 0.45]
More than four-year degree	0.13***	[0.05, 0.33]
Age group by education interaction		
(Age 25–29) and (BA or greater)	2.49*	[1.03, 5.98]
(Age 30–38) and (BA or greater)	4.33**	[1.84, 10.23]
(Age 39+) and (BA or greater)	4.02**	[1.61, 10.04]
Marital status (Ref: never married)		
Formerly married	1.49**	[1.18, 1.89]
Currently married	0.93	[0.82, 1.06]
Year (Ref: 2003)		
2004	0.78	[0.54, 1.13]
2005	0.94	[0.66, 1.35]
2006	1.29	[0.93, 1.79]
2007	1.33	[0.96, 1.83]
2008	1.54**	[1.12, 2.10]
2009	1.90***	[1.41, 2.57]
2010	1.84***	[1.36, 2.49]
2011	1.81***	[1.34, 2.45]
2012	2.30***	[1.72, 3.09]
2013	1.77***	[1.30, 2.41]
2014	1.94***	[1.43, 2.63]
2015	1.92***	[1.41, 2.62]
AIC	26164.44	
Observations	7,249,908	

NOTE: Calculation of Tjur's *D* statistic, which measures the difference between the mean predicted value of suicide for those who do and do not die of suicide within our analytic sample, yields a value of 0.00005. This suggests that while results show matchable factors that have large and significant associations with suicide risk, because suicide is a highly rare event in our data set, the overall ability of our model to discriminate between soldiers at risk of suicide is low.

***$p<0.001$, **$p<0.01$, *$p<0.05$.

(OR = 0.20; 95 percent CI = 0.09, 0.45), although the protective effects of higher education are greater for individuals under age 25 (OR = 0.13; 95 percent CI = 0.05, 0.33). Finally, formerly married individuals have nearly 50 percent greater odds of suicide relative to those who are currently married (OR = 1.49, 95 percent CI = 1.18, 1.89).

These results are consistent with those from prior research showing higher suicide rates among soldiers who are male; non-Hispanic white, Asian/Pacific Islander, or Native American (the latter of whom would be captured under "Other" in Table 3.3's specification); under age 21; and of lower educational status. While our results differ from prior evidence showing a significant negative effect of being married on suicide risk (Schoenbaum et al., 2014), this difference may be attributable to how our study operationalized marital status (i.e., we distinguish soldiers who were never married from those who were formerly married) as well as our use of multivariate models rather than the bivariate models commonly used in prior work. Indeed, in bivariate analyses (not shown), estimates from our sample show a significant negative association of being currently married with suicide risk. This suggests that differences in sociodemographic characteristics (e.g., age, education) across marital status categories may contribute to the observed negative association of being currently married with suicide risk in unadjusted models. For comparative purposes, exploratory bivariate analysis using our sample also showed a significant negative association of suicide risk with being currently married (OR = 0.86; 95 percent CI = 0.77, 0.96); however, once our models adjust for other demographic factors (see Table 3.3) the magnitude of this relationship was reduced and became insignificant.

Table 3.3 also shows clearly significant effects of time. As shown in prior research (Watkins et al., 2018; Nweke et al., 2015), suicide rates in the Army have trended upward since 2004, with a significant increase in risk relative to 2003 occurring in 2008 and persisting through 2015. Suicide rates peaked in 2012 with adjusted odds of suicide over twice as high relative to 2003 (OR = 2.30; 95 percent CI = 1.72, 3.09). In 2015, adjusted odds of suicide were nearly twice as great as in 2003.

As noted in Chapter Two, the rise in suicide rates has also been observed in the general U.S. population with the national suicide rate increasing from 10.5 to 13.0 per 100,000 (Curtin, Warner, and Hedegaard, 2016). Thus, our finding is consistent with more general trends.

Second-Stage Models of Unmatchable Army-Relevant Factors

We used the estimated coefficients from the final multivariate baseline model (Table 3.3) to estimate predicted log-odds of suicide for each person-year in the data. Then, to understand Army-specific factors associated with suicide risk above and beyond our set of matchable factors, we estimated logistic regressions with each unmatchable Army-specific factor included as a covariate of interest independently and the predicted log-odds from our baseline model entered as an offset (i.e., with coefficient constrained to one).

Table 3.4 presents results from each of these regressions, showing how unmatchable Army-specific factors and number of dependents relate to suicide risk after adjusting for the matchable baseline factor. Each column in Table 3.4 represents a separate regression with columns 1, 2, 3, and 4 examining the potential lingering associations between suicide rates and deployment history, foreign unit location, rank, and number of children, respectively, after controlling for the matchable factors.

The model that examines deployment history (column 1) suggests that those who have a history of deployment are at increased risk, after accounting for matchable baseline factors (OR

Table 3.4
Effect of Unmatchable Factors on Suicide Risk After Controlling for Matchable Characteristics

	Deployment Model (1)	Location Model (2)	Rank Model (3)	Children Model (4)
Deployed once	1.38*** [1.21,1.57]			
Deployed 2+	1.19** [1.04,1.35]			
Not in the U.S.		0.68*** [0.56,0.82]		
E3			1.19 [0.97,1.45]	
E4			1.29** [1.08,1.54]	
E5			1.27* [1.04,1.54]	
E6			1.01 [0.80,1.26]	
E7–E9			1.21 [0.96,1.53]	
O0–O2			0.99 [0.66,1.49]	
O3			1.15 [0.83,1.61]	
O4+			0.98 [0.67,1.42]	
WO			0.65 [0.40,1.05]	
One child				1.01 [0.87,1.17]
Two children				0.96 [0.82,1.13]
3+ children				1.01 [0.86,1.19]
Observations	7,249,908	7,249,908	7,249,908	7,249,908
AIC	26,087.8	26,091.3	26,103.9	26,114.1
AIC diff	–20.63	–17.14	–4.52	5.69

NOTE: Odds ratios and 95 percent confidence intervals from logistic regressions presented. AIC diff presents the difference in AIC statistics between the presented model and a model only including the offset and constant terms.

***P<0.001, **p<0.01, *p<0.05.

= 1.38 for one prior deployment versus never deployed; 95 percent CI = 1.21, 1.57 and OR = 1.19 for 2+ prior deployment versus never deployed; 95 percent CI = 1.04, 1.35). These findings, in turn, suggest that the difference between suicide rates in the Army and general populations will likely vary as a function of deployment history among soldiers. Army suicide rates will look highest relative to the matched general population when looking at those soldiers with prior deployments or when comparing years in which a higher proportion of the Army population had a deployment history. Appendix E explores how the comparisons between Army and general population suicide rates vary as a function of deployment history over time.

The model examining unit location (column 2) shows that, after controlling for baseline characteristics, there remains a significant and substantive relationship between suicide and foreign unit location, with soldiers in foreign locations having lower suicide rates than soldiers in the United States (OR = 0.68; 95 percent CI = 0.56, 0.82). Considering this finding, comparisons between the Army and general U.S. population should carefully consider whether or not to include soldiers serving overseas. Because we do not have comparable data on the suicide rate of members of the general U.S. population who live abroad, analyses comparing Army with general populations could exclude soldiers not in CONUS from the sample since including them may underestimate the difference in suicide rates between the domestically located Army and matched general population. However, the extent to which the significant difference in suicide risk between U.S.-located and foreign-located soldiers shown in Table 3.4 has practical implications for matched comparisons between the Army and general populations will depend on the proportion of the Army located overseas. Because foreign-located soldiers comprise a small proportion (13.4 percent) of our Army sample over the study period, excluding soldiers who are overseas does not substantially affect estimated suicide rates (see Appendix E).

In the model that examines rank (column 3), after adjusting for matchable factors, the rank variables are jointly significant ($\chi^2(9) = 20.9$, $p = 0.01$). Our results show that soldiers with E4 or E5 rank have significantly higher rates of suicide than other ranks (OR = 1.29 for E4 versus E0–E2; 95 percent CI = 1.08, 1.54; and OR = 1.27 for E5 versus E0–E2; 95 percent CI = 1.04, 1.54). There is also a trend toward soldiers in the E3 rank having higher rates of suicide even after adjusting for baseline factors including age (OR = 1.19 for E4 versus E0–E2; 95 percent CI = 0.97, 1.45), although this association is only marginally statistically significant ($p = 0.09$). Thus, it is likely that the observed difference in suicide risk between Army and similarly matched general populations will vary as a function of rank, with greater differences being observed among soldiers in the enlisted ranks of E3–E5 and smaller differences observed between officers and E1–E2 soldiers and similarly matched members of the general population.

The model that examines having children (column 4) shows that, after controlling for baseline characteristics, there is no significant relationship between the number of children and suicide risk among the Army sample.[2]

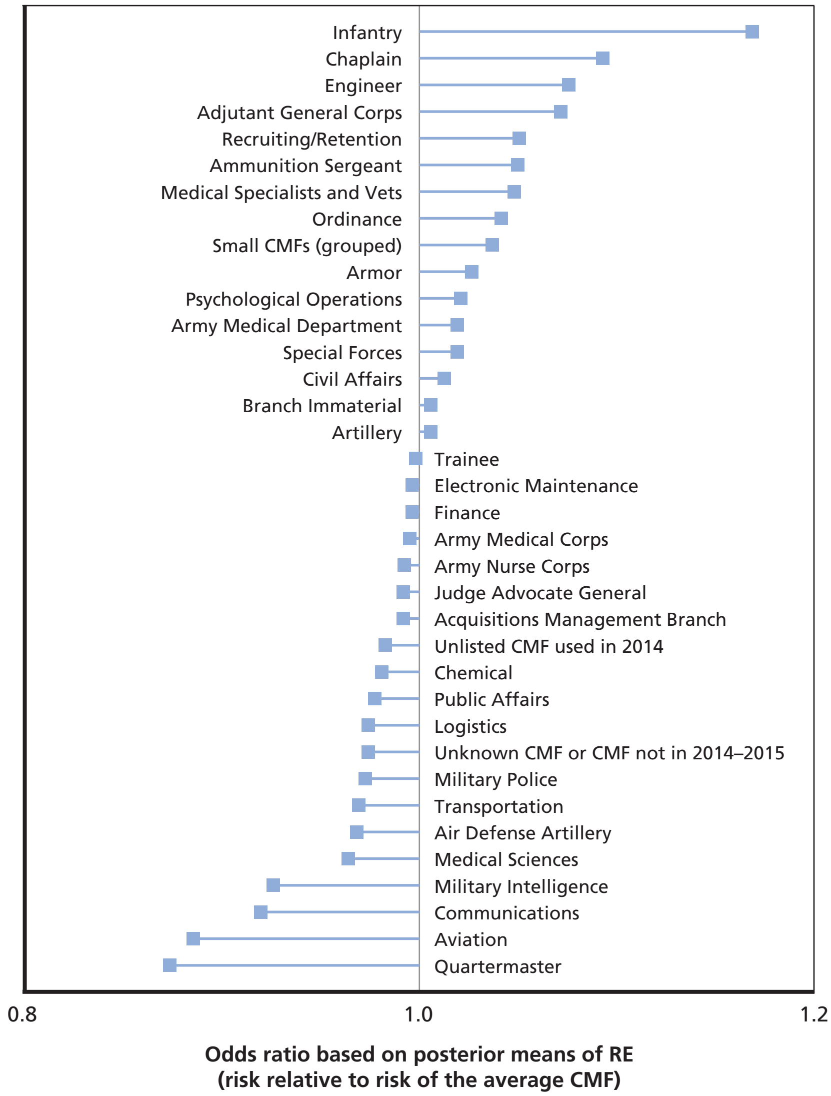
Finally, Figure 3.1 presents results from the regression model treating occupation (i.e., CMF category) as the unmatchable factor of interest. Each point on the graph (calculated as the posterior mean of the random effects) represents the estimated odds ratio for a given CMF category in relationship to suicide risk among the average CMF, and CMFs are ordered from highest to lowest risk. While results showed occupation to be significantly associated with suicide risk after controlling for matchable factors ($p = 0.002$), the magnitude of the associations for any given CMF is relatively small. Estimated odds ratios for the lowest risk category (Quartermaster) and highest risk category (Infantry) span 0.87 to 1.17.

Conclusions

Findings from this chapter confirm that age, gender, race/ethnicity, educational attainment, marital status, and year explain variation in suicide risk in the Army, highlighting the important role these matchable factors should play when trying to compare Army and general population suicide rates. Additionally, there are key unmatchable Army-relevant factors that explain

[2] Although the number of children does not seem to matter for the Army (after controlling for other factors), it may still matter for the general population, in which case we would still want to match on it to provide the most direct comparison between the two populations. Unfortunately, we do not have information on number of children in our civilian data on suicide victims. In this case, number of children is distinct from rank, deployment history, and foreign unit location, because it is not an Army-specific factor, but rather it is considered as part of unmatchable characteristics because the NVDRS civilian data set does not have this variable on civilians.

Figure 3.1
Estimated Associations of Career Management Fields with Suicide Risk



NOTE: Each point represents the estimated odds ratio for a given CMF category relative to suicide risk among the average CMF after controlling for matchable factors.

variation in suicide risk above and beyond these factors. These additional Army-relevant factors are deployment history, unit location, rank, and military occupation. We note that these variables are different from the unmatchable factors identified as part of our literature review (geography, parenthood, occupation, mental health, and firearm availability); however, we use these Army-relevant factors to inform the potential role omitted variables might have in putting the suicide rate comparisons reported in Chapter Five into perspective.

For matched comparisons, it is likely that the difference between the Army and general population suicide rate will vary as a function of these unmatchable factors, though it is unclear by how much. Thus, the results from our analyses on unmatchable Army-relevant factors act as a way to assess the potential impact of unobserved covariates in our comparisons between

Army and general population suicide rates presented in Chapter Five. The substantive implications of how not being able to adjust for unmatchable factors will impact differences in suicide rates between the Army and the general population will depend both on the factor's effect size associated with suicide risk as well as on the proportion of the Army exhibiting the significant factors. There is thus a need to consider both the strength of the association between a given unmatchable factor as well as its relevance for the Army. For example, excluding Army soldiers who are overseas from comparisons between the Army and general populations might increase the difference in suicide rates reported in typical Army versus general population analyses. However, given the relatively small proportion of soldiers located overseas during our study period (13 percent), excluding them from a matched comparison may actually do very little to alter estimated suicide rate differences. In contrast, excluding service members with a history of previous deployment (who comprise nearly 60 percent of our sample) may lead to more dramatic differences in suicide rate comparisons with the general population. Appendix E explores these issues in more detail.

CHAPTER FOUR

General Population Risk Factors

In this chapter, we describe our analyses of general population risk factors. These analyses served two primary objectives. First, RAND Arroyo Center aimed to identify the range of general population data sources that might be used to make suicide rate comparisons with the Army. Second, we aimed to use the general population data to understand (and confirm) the extent to which matchable factors identified in Chapter Two explain variation in suicide risk in our data set. In the case of our general population analyses, we did not have a set of unmatchable factors to consider; thus, we only present findings regarding the relationship between matchable factors and suicide risk.

Data Sources

Several databases are available to examine suicide risk in the general population (described in Appendix C). Our goal was to select one that was representative of the U.S. population, or a subset of it, and that included an expanded set of factors with which to match to the Army sample. We needed to have data on suicide cases in the general population along with information on the key factors identified in Chapter Two as explaining variation in suicide risk.

In terms of suicides themselves, the NVDRS is the only available state-based reporting system that pools data from multiple sources into a usable, publicly available database on violent deaths. The sources providing details on violent deaths include state and local medical examiner, coroner, law enforcement, toxicology, and vital statistics records. As such, the NVDRS is currently the richest source of information on general population suicides. It provides detailed information about decedents; their cause of death; their family and life circumstances; and information about their education, occupation, age, and other demographic characteristics. This comprehensive set of factors are not all available in other data on suicides, such as the National Vital Statistics System.

However, a key limitation of the NVDRS is that it contains only suicide information on a subset of states. Only a handful of states have participated in the NVDRS collection over the study period; as a result, NVDRS currently contains data on just 27 states, excluding some of the most populous states in the country. In spite of this limitation, we opted to use the NVDRS because of the rich covariate information given on each suicide case.

We note that due to changes in the states represented in NVDRS over the study period, there are some large geographic shifts in our general population sample that will contribute to the changes observed in suicide rates over time; thus, we will have both time trends and changes in populations to consider when discussing our findings. Future plans include expanding the NVDRS to 40 states, which will allow future work replicating our methods to have greater generalizability than the results we report.

The NVDRS contains only detailed data on suicides. To establish whether marital status, education level, or other key factors are associated with suicide risk among the general U.S. population, we also needed information on whether those characteristics are over- or underrepresented among the suicide cases as compared with the more general population from which the suicide cases arose. Therefore, we merged the suicide cases from the NVDRS with the CPS, a general population database that contains representative data of the general populations living in the states included in the NVDRS in each year. The CPS is a monthly survey of about 60,000 U.S. households conducted by the U.S. Census Bureau for the Bureau of Labor Statistics (BLS). The BLS uses the CPS data to publish monthly reports on the unemployment rate and the numbers of employed and unemployed people in the United States. The CPS is a nationally representative and state-by-state representative survey providing high-quality information about the characteristics of the general U.S. population overall and the population of the NVDRS states. Thus, the combined data sets provide us with usable information on marriage and educational categories for those who died by suicide and for those who did not, a necessary condition for estimating risk.

The specific CPS data we used was the March Annual Social and Economic Supplement (ASEC) sample of the CPS for each year from 2003 through 2015. The March ASEC sample was chosen due to our initial focus on occupational variation in suicide risk and the detailed occupation-related information available in the ASEC supplement to the CPS.

The ASEC of the CPS is the nation's source of official estimates of poverty levels and is widely used for measures of income and employment. It provides annual estimates based on a survey of more than 75,000 households. The survey contains detailed questions covering social and economic characteristics, including income and job type. The survey respondents are assigned sampling weights so that the sample provides national and state-representative rates for the key quantities measured in the survey. These weights are augmented in our data fusion process to ensure continued representative after we fuse NVDRS to the CPS data in each calendar year.

Because suicide risk and protective factors may vary geographically, we subset the CPS data to only NVDRS states in each year and then imputed suicide cases within the CPS by state and year. The imputation is assigning an unknown value to the CPS data (namely, an indicator for whether the CPS case experienced a suicide). The value is based on how similar the CPS case is to suicide cases from the NVDRS on the observed factors we have available in both databases, such as age, gender, race/ethnicity, education, marital status, and geography. As such, our findings are only generalizable to the subset of the general U.S. population living in NVDRS states.

Because we know geographic differences exist in state suicide rates, suicides in the NVDRS are likely to differ in systematic ways from suicides nationally, so the NVDRS data set is not representative of the entire U.S. population. As a result, a comparison of Army rates with a NVDRS-CPS subset sample drawn from just the states participating in the NVDRS could not be used to understand how risk in the Army differs from the risk typically experienced nationally.

Constructing the Fused NVDRS-CPS Data Set

Given the independent pieces of information provided by the NVDRS and CPS, it was critical to design a method for merging the two databases. We implemented a novel data fusion strategy to obtain representative data from a subset of the general U.S. population that merged data from the NVDRS with data from the CPS. Data fusion is a process of integrating data from multiple sources to produce new databases that can be used to provide more useful and accurate information than available in any single data source (Das, 2008). In our data fusion efforts, our goal was to accurately identify individuals in the CPS who are comparable with suicide cases from the NVDRS on as many overlapping characteristics as possible between the two data sets in order to impute suicide information into the CPS. We then followed the procedures described below to fuse the two databases over a 13-year period between 2003 and 2015.

Figure 4.1 highlights how merging the Army and general population data differed. For the Army, we could easily merge suicide information into the larger Army database using an identifier that directly links suicide cases with their demographic and military data. For the NVDRS-CPS sample, data fusion was needed to robustly merge two disparate data sources.

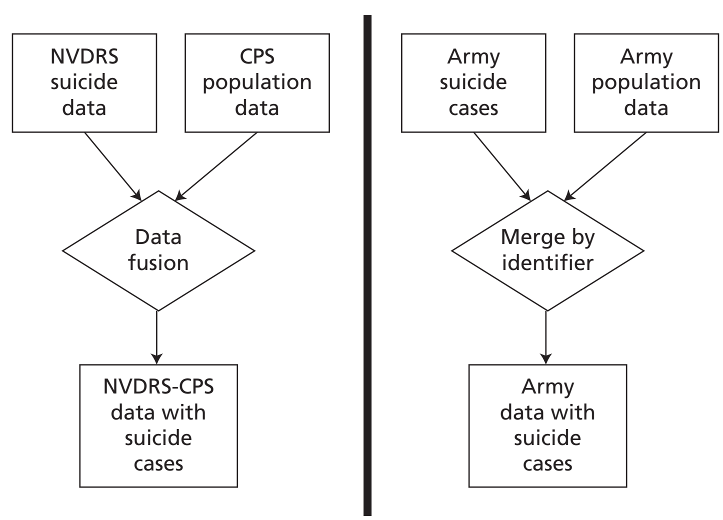
Data Restrictions

Both the NVDRS and CPS data files were subset to include only individuals aged 18 or older. Within each calendar year from 2003 through 2015, the CPS and NVDRS data were further subset to only the NVDRS states for which that entire year of data is available. Postfusion data were then restricted to ages 18 to 75. Our final sample thus contained 810,584 person-year records that, once weighted using our adjusted CPS weights (see details below), represent a total population size of 782,087,485 person-years and 130,040 suicides across all years.

Missing Data

Following the initial data cleaning, our data had very limited missing values for demographic variables (range = 0.0 percent to 2.64 percent). As such, we handled missing data by coding those individuals with missing values as part of the same group as those in the most frequently observed group from our data. Thus, for education, individuals with missing data were include

Figure 4.1
Data Structure for NVDRS-CPS and Army Databases



as part of the "less than BA" group; for race/ethnicity, the non-Hispanic white group; for marriage, the currently married group; and for gender, as part of the male group; regarding age, it was set equal to the median age of 26.

Identifying Overlapping Characteristics

We identified 11 characteristics that overlapped between the CPS and NVDRS: age, gender, race, Hispanic ethnicity, marital status, education level, veteran status, whether the person was born in the United States, state, census region, and whether the person lived in a rural or urban area. In some cases, the NVDRS and CPS coded variables differently. In order to fuse the two databases, the variables had to be harmonized—that is, we had to make the variables have the same format and values so they could be more easily merged. In most cases, finer grained categories from one data set had to be collapsed to the broader categories in the other. A full crosswalk of how variables were harmonized is presented in Appendix D.

Similarity Scores

Our goal was to identify individuals in the CPS who looked identical to the suicide cases in the NVDRS within the same year so that we could then identify who in the CPS should be imputed to have experienced a suicide. For example, looking only at age and gender: consider a suicide case from the NVDRS who is age 24 and male as well as two CPS respondents. Respondent 1 is 24 and male, and respondent 2 is 24 and female. Our suicide case would be an exact match on age and gender with respondent 1 but not with respondent 2. In this simple example, we would want to impute the CPS respondent 1 to be counted as a suicide during our data fusion process. To do this explicitly, we add a duplicate (new) copy of CPS respondent 1 (24-year-old male) to the original CPS file and flag it as a suicide case. This "imputed" suicide case would get a weight of one and the original CPS respondent would have its CPS population weight reduced by one. This is because the CPS population weight represents the total number of individuals in the nation with the characteristics of the given CPS respondent, and we have removed one of those individuals to now denote an individual who experienced suicide.

In implementing the data fusion process, we did not expect that all NVDRS suicide cases would have a perfect match on all variables to cases in the CPS data. Thus, we needed to develop an approach to fusing the data that allowed for matching when there was an imperfect match. To do this, we calculated "similarity scores" for cases within the CPS data that essentially produces a score for how similar each CPS case is to each suicide case in the NVDRS. Thus, each CPS respondent has multiple similarity scores mapping them to all the NVDRS suicides with which it had any overlapping characteristics. Each similarity score computed how similar the given CPS respondent was to the NVDRS suicide on all the covariates used in the data fusion process. The final set of variables used in computing the similarity scores, their values, and the weight they contributed to the similarity score is presented in Table 4.1.

To illustrate how it works when there are no perfect matches, let us assume that we have three variables on which we are matching (age, gender, and education). We have a CPS respondent who is a 23-year-old male with a bachelor's degree and two NVDRS suicide cases to which the CPS respondent can be matched. Suicide case A is a 28-year-old male with a bachelor's degree, and suicide case B is a 24-year-old male with an associate's degree. The similarity scores for this CPS respondent would be 2 to suicide case A and 1 to suicide case B, respectively (with 3 being a perfect score), given the extent to which the characteristics of this CPS respondent overlap with the suicide cases from the NVDRS. This scoring process generalizes

Table 4.1
Data Fusion Variables

Variable	Weight	Values
Age, bin 1a	0.25	[18, 22), [22, 26), [26, 30), …, [78, 82), [82+]
Age, bin 1b	0.25	[18, 20), [20, 24), [24, 28), [28, 32), …, [76, 80), [80+]
Age, bin 2a	0.25	[18, 30), [30, 42), [42, 54), [54, 66), [66, 78), [78+]
Age, bin 2b	0.25	[18, 24), [24, 36), [36, 48), [48, 60), [60, 72), [72, 84), [84+]
Sex	1	Female, Male
Race	1	White, Black, Asian/Pacific Islander, American Indian, Two or more
Ethnicity	1	Hispanic, Not Hispanic
Education bin 1c	0.5	≤ 8th grade, 9th–12th grade, high school or GED, some college, associate's, bachelor's, master's, doctorate
Education bin 2	0.5	No diploma, high school or GED, associate's or bachelor's, master's or higher
Marital status	1	Never married/single, married, divorced or separated, widowed
Veteran status	1	Yes, no
Born in U.S.	1	Yes, no
State	0.5	Alaska, Arizona, Colorado, Connecticut, Georgia, Hawaii, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia, Wisconsin*
Census region	0.5	East North Central, East South Central, Middle Atlantic, Mountain, New England, Pacific, South Atlantic, West North Central, West South Central
Flag for urban/rural	1	Urban, rural

* Because matching was done within a year, only states that were in the NVDRS for a given year were used in the match.

readily to allow for matching on all the variables listed in Table 4.1 and to allow for some additional nuisances in matching on age and education. Age and education were operationalized in multiple ways to compensate for the somewhat arbitrary bins we created for these variables. Additionally, state and census region were only given a weight of 0.5 since they are nested geographical variables and giving each a full weight of 1 would have overweighted the role of geography in matching.

In our matches, the highest possible similarity score is a 10, though we consider similarity scores over 8 as high-quality matches. Table 4.2 presents summary statistics of the quality of the fusion process. Over 95 percent of the suicide cases matched with a similarity score of 8.00 or higher, and 20.1 percent of all suicides received a perfect match on shared variables. The lowest similarity score for any match was 4.00. Fewer than 5 percent of suicide cases were matched with a similarity score under 8.00; most scores fell between 7.00 and 7.99. Each suicide case was matched with 4.0 CPS respondents on average. Poorer matches tended to include more CPS respondents, which will introduce some random error into estimates for those variables on which they did not match. However, with the large majority of CPS respondents matching at 8.00 or higher, the impact of this error should be small.

Table 4.2
Summary of Similarity Scores for Fused NVDRS-CPS Data

Similarity Score	*N* Suicide Cases	Percent of Suicide Cases	*N* Matches	Average Matches per Suicide Case
4–4.99	12	< 0.01%	210	17.5
5–5.99	155	0.1%	1,939	12.5
6–6.99	846	0.6%	8,370	9.9
7–7.99	4,551	3.3%	20,397	4.5
8–8.99	32,007	23.0%	139,228	4.4
9–9.99	73,660	52.9%	313,215	4.3
10	28,064	20.1%	78,323	2.8

Data Fusion

As noted, for each NVDRS suicide case, the CPS respondent or respondents with the highest similarity score were duplicated in the original CPS file and tagged as a suicide case. A single NVDRS suicide could match with one or multiple CPS respondents due to the possibility of multiple CPS respondents sharing the highest similarity score with a given NVDRS suicide. Also, a particular CPS respondent may be duplicated more than once if it has the highest similarity score for more than one NVDRS suicide case. In the next section, we describe how we adjust the CPS sample weights to compensate for our creating these duplicates.

Weighting

As described above, the CPS is a sample, and each observation in the CPS has a weight that can be used to make the overall CPS sample representative of the nation or a particular state. Thus, when suicide cases (duplicates) are added to the CPS sample using our data fusion, we needed to update and adjust the sampling weights to ensure our fused data maintained the overall representativeness of the CPS sample. That is, we wanted to maintain the same weighted population size and characteristics.

To compute the needed adjusted weight, we implemented the following steps. First, we created a weight for the newly added duplicate CPS respondents that are used to represent our suicide cases and correspondingly reduced the weight of the original CPS respondents to ensure that we maintained the right overall representativeness of those respondents. More specifically, we assigned each of the duplicated respondents tagged as a suicide a sample weight that equals one divided by the total number of matches for that suicide case (thus the sum of the weights for the duplicate respondents that matched to that one suicide will sum to one). If the CPS respondent is matched across more than one suicide case, the final weight for that CPS respondent is the sum of these weights. Then, in the second step, all original CPS respondents that were duplicated as suicides have their original weights reduced by the weight of their suicide duplicate.

To illustrate our weighting calculations, consider the following example continued from above: suppose each of the CPS cases had a population weight of ten. Since cases (1) and (2) matched with the first suicide, they would be duplicated and tagged as suicides with a weight of 0.5 (one divided by the number of matches). However, they also matched with suicide case

(2) where there were three total matches; therefore, their weights receive an additional 0.33 resulting in suicide cases for (1) and (2) with weights of 0.83 each. Then the original case (1) and (2) would have their weight reduced by 0.83 each to 9.17 to keep the weighted population size and characteristics consistent.

Suicide Rate Estimates and Comparison with CDC WONDER

Table 4.3 shows how the fused NVDRS-CPS data set compares with CDC WONDER suicide rates for the subset of NVDRS states and for the total of all states. As expected, the suicide rates from our fused data set align closely to the CDC WONDER rates for the subsample of NVDRS states, though the fused data set yields rates consistently higher by 0.1 to 1.0 suicide per year per 100,000. These types of differences were also found in related studies, with NVDRS tending to provide slightly higher estimates of suicide rates than CDC WONDER, which is based on data from the National Vital Statistics System (NVSS) (CDC, undated-b; CDC, 2005). This small discrepancy between NVDRS and NVSS occurs because NVSS restricts reporting to deaths of legal residents of the selected states while NVDRSs count all suicides that occur within the state regardless of legal residence (Regoeczi and Banks, 2014). In addition, another study showed that NVSS tends to underestimate violent deaths in general, which likely contributes to its slightly lower suicide rates as compared with our NVDRS-CPS rates (Brown et al., 2013).

When comparing the two CDC WONDER rates between the subset of NVDRS states and the overall United States, the NVDRS states have slightly higher suicide rates on average (approximately –1 suicide on average per 100,000 people across years).

Table 4.3
Comparison of Suicide Rates in NVDRS-CPS Merged Data on Subset of NVDRS States Versus National Rates from CDC WONDER

Year	CDC WONDER Rate per 100,000: All United States	CDC WONDER Rate per 100,000: NVDRS States	NVDRS-CPS Merge per 100,000
2003	13.9	12.3	13.0
2004	14.1	14.0	14.9
2005	14.0	14.3	15.1
2006	14.3	14.8	15.6
2007	14.7	15.2	15.9
2008	15.1	15.3	16.0
2009	15.3	15.9	16.7
2010	15.9	16.5	16.9
2011	16.1	16.5	17.2
2012	16.4	17.0	17.6
2013	16.3	17.0	17.9
2014	16.8	17.3	18.0
2015	17.2	17.2	17.4

Analysis of Suicide Risk Factors

The objective for our regression modeling within the general population is to understand how the matchable factors presented in Chapter Two relate to suicide risk within the general population. These analyses provide suggestive evidence for how matching (or failing to match) on these factors may affect suicide rate comparisons between the Army and general population samples. For example, if marital status is associated with suicide in the general population, then because marriage rates differ between the Army and general populations, Army-general population comparisons that adjust for only age and gender will fail to adequately adjust for this factor and misrepresent the true differences between suicide rates in the two populations. In addition, the modeling serves as a validity check of our CPS-NVDRS merging process by ensuring that the relationships estimated in our data set reflect the known relationships between suicide risk and these factors, as demonstrated in prior literature (see Chapter Two).

Our approach to modeling suicide risk in the general population is analogous to step 1—the baseline analysis—described in Chapter Three for the Army analysis.

As with the Army analyses, we use person-year covariates as the primary predictors of suicide risk in the general population. We began by fitting bivariate models for each of the covariates used in the fusion process, then estimated models examining only age, gender, and race/ethnicity, as has been common in prior work on this topic. In addition, we considered one set of pairwise interactions (age group interacted with gender) given gender differences in the age profile of suicide risk that have been documented in previous research (NIMH, 2016). Based on these initial results, education level and marital status showed meaningful variation and were included in a full model built on the age by gender and race specification documented in prior literature. We also included fixed effects for year.

The variables examined across the bivariate and multivariate estimates include

- **gender**: male; female
- **age**: age groups 18–27; 28–37; 38–47; 48–57; and 58+
- **race/ethnicity**: American Indian, non-Hispanic Asian or Pacific Islander, non-Hispanic African American, Hispanic, two or more races or other
- **education level**: less than HS; HS diploma or equivalent; associate's or BA; master's or higher
- **marital status**: married; never married; divorced or separated; widowed
- **age by gender interactions**: gender with age groups 28–37, 38–47, 48–57, and 58–75
- **year**: dummy variables for each year 2003 to 2015.

Results

Population Characteristics

Table 4.4 shows the demographic characteristics for our NVDRS-CPS sample from 2003 to 2015. Due to changes in the states represented in NVDRS over the study period, there are some large geographic shifts in our NVDRS-CPS sample. There is no West North Central region representation until 2015, when Kansas and Minnesota joined the NVDRS. The South Atlantic region comprises the largest proportion of the sample across all years, although this proportion decreases from 45 percent–50 percent to 40 percent of the sample to 26.8 percent of the sample in 2015. Conversely, the East North Central region is poorly represented early on but comprises 20 percent of the sample in 2011 when Ohio joins the NVDRS.

Table 4.4
Descriptive Statistics for Weighted NVDRS-CPS Sample Characteristics from 2003 to 2015

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
N	25,023,339	47,057,893	53,350,188	54,251,839	55,039,133	55,833,248	56,295,922	57,065,684	65,555,681	66,005,123	66,321,331	74,198,645	106,089,459
Suicides/100,000	13	14.9	15.1	15.6	15.9	16	16.7	16.9	17.2	17.6	17.9	18	17.4
Gender													
Female	51.3%	51.1%	51.1%	51.2%	51.1%	51.0%	51.1%	51.3%	51.4%	51.6%	51.4%	51.0%	51.2%
Race/ethnicity													
White	71.6%	71.7%	71.9%	71.5%	70.9%	70.8%	70.6%	69.5%	71.0%	70.6%	70.0%	70.0%	68.0%
Black	15.0%	14.9%	13.9%	14.0%	14.2%	14.4%	14.4%	14.4%	13.9%	14.1%	14.0%	14.2%	12.8%
Hispanic	7.5%	7.9%	8.5%	8.9%	8.7%	8.8%	8.7%	9.4%	8.8%	8.8%	9.2%	8.8%	11.1%
Asian/Pacific Islander	4.6%	3.6%	3.7%	3.4%	3.9%	3.9%	4.0%	4.2%	4.0%	4.0%	4.3%	4.6%	5.8%
American Indian	0.6%	0.8%	0.9%	0.9%	1.1%	1.1%	1.2%	1.2%	1.1%	1.0%	1.1%	1.1%	1.0%
Two or more or other	0.9%	1.1%	1.1%	1.2%	1.1%	1.1%	1.1%	1.3%	1.3%	1.5%	1.4%	1.3%	1.3%
Mean age	43	42.8	42.8	42.9	43.2	43.2	43.4	43.7	43.9	44	44.2	44.5	44.6
Education level													
<HS	12.9%	13.6%	13.5%	13.3%	13.1%	12.1%	11.9%	11.5%	11.0%	10.8%	10.8%	10.3%	10.3%
HS/GED	48.8%	50.7%	51.0%	50.8%	49.4%	49.3%	49.5%	49.9%	50.5%	50.2%	49.4%	49.2%	47.8%
Associate's or bachelor's	27.5%	26.3%	26.2%	26.3%	27.6%	28.0%	27.8%	28.1%	27.9%	28.4%	28.5%	29.0%	30.1%
Master's or more	10.8%	9.4%	9.3%	9.7%	9.9%	10.6%	10.8%	10.5%	10.5%	10.6%	11.2%	11.5%	11.8%
Marital status													
Married	57.2%	57.6%	57.6%	57.1%	57.8%	56.6%	56.7%	55.9%	55.6%	54.7%	54.1%	54.5%	53.4%
Never	26.9%	26.4%	26.1%	26.3%	26.0%	27.1%	27.3%	27.4%	28.0%	28.6%	28.9%	29.2%	30.5%
Divorced/separated	12.2%	12.6%	12.8%	13.1%	13.1%	13.0%	12.9%	13.2%	13.2%	13.4%	13.7%	13.1%	12.8%
Widowed	3.7%	3.4%	3.6%	3.5%	3.2%	3.2%	3.1%	3.5%	3.3%	3.4%	3.3%	3.2%	3.3%

Table 4.4—Continued

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Veteran status													
Veteran	10.3%	9.7%	9.6%	9.3%	9.0%	8.8%	8.5%	8.2%	8.1%	8.0%	7.6%	7.3%	7.0%
Birthplace													
Born in U.S.	85.3%	87.2%	88.0%	87.4%	86.7%	86.7%	87.0%	86.3%	87.3%	87.3%	87.0%	86.8%	84.5%
Census region													
East North Central	0.0%	8.1%	7.2%	7.0%	7.0%	6.9%	7.0%	6.9%	18.2%	18.3%	18.1%	25.6%	18.4%
East South Central	0.0%	0.0%	5.4%	5.3%	5.3%	5.3%	5.4%	5.4%	4.7%	5.7%	4.7%	4.3%	2.9%
Middle Atlantic	23.6%	1.5%	11.2%	11.2%	11.0%	10.7%	10.7%	10.8%	9.5%	9.3%	9.3%	8.5%	19.5%
Mountain	0.0%	6.7%	11.3%	11.7%	11.6%	11.8%	11.8%	11.6%	10.3%	10.2%	10.5%	9.5%	11.1%
New England	18.3%	11.0%	9.8%	9.6%	9.5%	9.4%	9.4%	9.5%	8.3%	8.3%	8.1%	7.4%	10.0%
Pacific	11.7%	6.3%	5.6%	5.5%	5.6%	5.7%	5.7%	5.6%	4.9%	5.0%	4.9%	4.5%	4.0%
South Atlantic	46.5%	50.3%	45.0%	45.2%	45.7%	45.9%	45.7%	45.8%	40.3%	40.4%	40.5%	36.8%	26.2%
West North Central	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	5.5%
West South Central	0.0%	5.0%	4.5%	4.5%	4.4%	4.3%	4.3%	4.4%	3.9%	4.0%	3.8%	3.5%	2.4%
Rural proxy													
Rural	55.8%	60.3%	64.3%	65.1%	65.1%	64.8%	64.8%	64.7%	62.9%	62.8%	62.7%	59.2%	53.1%

The size of the population in our sample increases over the entire period with large jumps occurring in 2004, 2011, 2014, and 2015—the years in which NVDRS added states. The suicide rate increases over time from 2003 through 2014, and then in 2015 the rate drops. Gender composition is consistent over time (around 51 percent female), but the average age increases from 43 to 44.6. Racial composition trends are also consistent over the 14-year period, though the percentage of non-Hispanic whites decreases from 71.6 percent in 2003 to 68.0 percent in 2015 while the percentage of Hispanics increases from 7.5 percent to 11.1 percent. Smaller shifts are observed for percentage of African Americans (decreases) and percentage of Asians (increases). There is little change in the percentage of American Indians and percentages of two or more or other race/ethnicity. Across all years, high school diploma or equivalent is the largest educational attainment group (range = 47.8 percent–50.8 percent), followed by those having an associate's or bachelor's degree (range = 26.2 percent–30.1 percent), those without a high school credential (range = 10.3 percent–13.6 percent), and those having a master's or higher (range = 9.3 percent–11.8 percent). A large majority of the NVDRS-CPS sample are married (range = 53.4 percent–57.6 percent) with never married/single representing above a quarter of the sample (range = 26.0 percent–30.5 percent). Around 12 percent are divorced/separated. There are few widowed persons likely due to the age cutoff at 75. Veteran status sees a steady decrease over the time period from around 10 percent to around 7 percent. The proportion of the NVDRS-CPS sample born in the United States remains consistent at about 85 percent.

Baseline Model for Matchable Factors

Table 4.5 presents findings from our baseline model that includes the matchable factors described in Chapter Two. Consistent with the literature, females have much lower odds of suicide (OR = 0.23, CI = 0.23, 0.24). For men, all age groups have higher odds of suicide compared with age 18–27. For women, the odds of suicide are higher for age groups 28–37 (OR = 1.23, CI = 1.17, 1.29), 38–47 (OR = 1.40, CI = 1.34, 1.47), and 48–57 (OR = 1.40, CI = 1.33, 1.46) compared with the age group 18–27, but the odds of suicide are lower for the age group 58–75 (OR = 0.92, CI = 0.88, 0.96) compared with ages 18–27. Black, Hispanic, and Asian/Pacific Islander race/ethnicity have very low odds of suicide compared to whites while American Indian and two or more races/ethnicities or other have moderately lower odds of suicide than whites.

Our results on education generally align with prior research. Prior work has found that suicide rates are highest among those with a high school degree and lowest among those with at least some college. Specifically, we find that individuals with higher levels of education have lower odds of suicide compared with individuals with a HS or equivalent diploma. Individuals with some college credit have an OR = 0.76 (CI = 0.75,0.78), associate's or occupational/vocational programing have an OR = 0.75 (CI = 0.72,0.77), bachelor's degrees have an OR = 0.66 (CI = 0.65, 0.68), master's or professional school degrees have an OR = 0.61 (95 CI = 0.60,0.62) and doctorate degrees have an OR = 0.82 (CI = 0.76, 0.87) relative to the group with a HS or equivalent diploma. Recent work has also shown that suicide rates are higher among those with only a HS diploma or equivalent compared with those with less than a HS diploma or the equivalent. Our results indicate that this is true for individuals who ended schooling in eighth grade or earlier (OR 0.78; CI = 0.75,0.82). However, we also show that individuals whose education ended between ninth and 12th grade have higher odds of suicide (OR = 1.16, CI = 1.131, 1.194) compared with those who have a HS or equivalent diploma is different from our related work. Our different findings for this particular subgroup (9th to 12th grade) could be a result of some coding differences or it could also be a result of our sample lacking nationally

Table 4.5
Baseline Multivariate Logistic Model of Association Between Suicide and Matchable Characteristics for NVDRS-CPS Sample

	OR	95% Confidence Interval (CI)
Gender (Ref: male)		
Female	0.234***	[0.227, 0.242]
Age group (Ref: age 18–27)		
Age 28–37	1.368***	[1.327, 1.410]
Age 38–47	1.585***	[1.536, 1.635]
Age 48–57	1.682***	[1.629, 1.736]
Age 58+	1.476***	[1.429, 1.524]
Gender by age group interaction		
(Female) and (Age 28–37)	1.229***	[1.172, 1.289]
(Female) and (Age 38–47)	1.401***	[1.339, 1.465]
(Female) and (Age 48–57)	1.395***	[1.333, 1.460]
(Female) and (Age 58+)	0.918***	[0.876, 0.962]
Race/ethnicity (Ref: non-Hispanic white)		
Black	0.336***	[0.328, 0.344]
Hispanic	0.383***	[0.372, 0.394]
Asian/Pacific Islander	0.388***	[0.372, 0.405]
American Indian	0.822***	[0.768, 0.880]
Two or more races or other	0.715***	[0.670, 0.763]
Education level (Ref: HS or equivalent)		
<= 8th grade	0.784***	[0.749, 0.820]
9th–12th grade, no diploma	1.162***	[1.131, 1.194]
Some college credit	0.763***	[0.747, 0.779]
Associate's or occupational/vocational program	0.745***	[0.724, 0.766]
Bachelor's degree	0.661***	[0.647, 0.676]
Master's or professional school degree	0.614***	[0.597, 0.632]
Doctorate degree	0.815***	[0.764, 0.868]
Marital status (Ref: married)		
Never married/single	2.289***	[2.242, 2.336]
Divorced/separated	3.053***	[2.992, 3.115]
Widowed	2.011***	[1.927, 2.098]

Table 4.5—Continued

	OR	95% Confidence Interval (CI)
Year (Ref: 2003)		
2004	1.135***	[1.081, 1.191]
2005	1.151***	[1.097, 1.207]
2006	1.180***	[1.125, 1.237]
2007	1.214***	[1.157, 1.273]
2008	1.226***	[1.169, 1.286]
2009	1.282***	[1.223, 1.344]
2010	1.306***	[1.246, 1.369]
2011	1.308***	[1.249, 1.370]
2012	1.336***	[1.276, 1.400]
2013	1.353***	[1.292, 1.417]
2014	1.366***	[1.303, 1.433]
2015	1.343***	[1.286, 1.404]
AIC	2,426,202	
Observations	782,087,485	

NOTE: Calculation of Tjur's *D* statistic, which measures the difference between the mean predicted value of suicide for those who do and do not die of suicide within our analytic sample, yields a value of 0.0001592. This suggests that while results show matchable factors that have large and significant associations with suicide risk, because suicide is a highly rare event in our data set, the overall ability of our model to discriminate between soldiers at risk of suicide is low.

***$p<0.001$, **$p<0.01$, *$p<0.05$.

representative data on the general population from many states, including the most populous: California, Texas, Florida, and New York.

Marital status appears to matter greatly and in the same way that prior literature has reported. Divorced or separated individuals have very high odds of suicide (OR = 3.05, CI = 3.00, 3.12) compared with married individuals. Never married and widowed persons also have higher odds (OR = 2.29, CI = 2.24, 2.34; and OR = 2.01, CI = 1.93, 2.10, respectively) compared with married persons. Finally, we again see strong effects of time with suicide rates trending upward since 2003 with the exception of 2014 to 2015.

Thus, although this general population is not nationally representative, we find the results of our analysis to reflect the findings of prior literature examining the relationship between these factors and suicide. The exception to this pattern is educational attainment.

Conclusions

Findings from this chapter confirm that age, gender, race/ethnicity, educational attainment, marital status, and time explain variation in suicide risk in the general population, highlighting the important role these matchable factors should play when trying to match the general population to the Army in order to compare suicide rates. Additionally, the chapter highlights our novel data fusion strategy for calculating NVDRS-CPS sample suicide rates by combining two general population data sources (the NVDRS and CPS). It is important to note that while the NVDRS contains rich information on suicide cases, it represents only a subset of states. Thus, our results and the comparisons we make with the Army presented in the next chapter are limited to only a subset of states and not to the nation. However, as the NVDRS expands to include more states, the same methods could be applied to make more nationally representative comparisons.

The CPS was chosen because it provides a representative sample of the general U.S. population and has high-quality data on general population occupations. Ultimately, we could not explore variation in suicide risk by occupation due to coding issues in the death records. In Appendix A, we describe our efforts to better understand how occupation and industry are coded in federal data systems and highlight the need for a collaboration between the Army, U.S. Census Bureau, the CDC, and Department of Labor to prioritize more accurate coding of death records.

CHAPTER FIVE

Matching the Army to a Comparable Subset of the General U.S. Population

Having assembled and analyzed the Army and NVDRS-CPS data separately, our final task was to understand the impact of using different sets of factors for comparing the general population suicide rate (as represented by the subset of the U.S. population represented in the states included in the NVDRS) with the Army suicide.[1] Specifically, we developed two sets of matching factors. The first set included commonly used demographic factors (i.e., age and gender) to match the NVDRS-CPS sample to the Army sample between 2003 and 2015. The second included the expanded set of matchable risk factors identified in Chapter Two (i.e., age, gender, race/ethnicity, education, and marital status). Both sets of factors also included time by matching the two populations within each calendar year.

In all analyses, we created weights that make NVDRS-CPS sample look like the Army sample in each calendar year. This is an important choice in our standardization. In principle, one could standardize in either direction—that is, reweight the Army to match the general population for a particular set of matching factors. In practice, however, it is preferable to weight the general population sample to match the Army for several reasons. For one, matching the Army to look like the general population would not work when trying to match on factors like age, because parts of the general population age distribution (namely, the oldest age groups) do not exist in the Army at all or in very small sample sizes. For another, it is substantively of interest to understand the impact of serving in the Army for soldiers like those who would serve. To answer this question, one needs to know what the suicide risk for soldiers would have been had they not joined the Army, which is necessarily unobservable but can be approximated if we can compare soldiers with members of the general U.S. population who did not join the Army but who otherwise look like Army soldiers after weighting.

Identifying Appropriate Statistical Strategies for Matching

As part of our work to identify statistical strategies for matching the Army to a comparable general U.S. population, we performed a high-level literature review of the various methods

[1] It is important to note that the subset of NVDRS states have slightly higher suicide rates than the general U.S. population, as presented in Chapter Four. It is unclear whether these higher rates are consistently observed within and across our matchable factors. As such, we cannot be fully sure how different the comparisons presented in this chapter would be if the Army suicide rate was compared with all states. In an ideal setting, we could determine the appropriate offset that would be needed to ensure our weighted civilian suicide rates have adequately adjusted its elevated suicide rate relative to the general population, but it is impossible to correctly compute this offset without knowing suicide rates for the full distribution of our matchable factors among the general U.S. population.

used to match one population to another. We identified three candidate methods—standardization, matching, and weighting—and ultimately selected weighting. While standardization is attractive because it does not require individual-level data, the approach tends to fall apart when used to match populations on multiple factors simultaneously (Watkins et al., 2018; Reimann and Mazuchowski, 2018). This often results in having to drop cases in one or both populations or having to greatly simplify the factors used to stratify. Similarly, matching is a useful way to ensure that two samples look alike on key characteristics like age and gender, but Army cases that do not have good general population matches would get dropped from the analysis and, in turn, would shift the underlying dynamic of a matching process and restrict how results and findings can be generalized.

The purpose of weighting is to create two groups that, when weighted, look "well balanced" or comparable on the range of characteristics that are used in the analysis. The goal is to obtain populations with the same distributions of the key factors that are associated with suicide (in our case, matchable factors). These similarities in turn ensure reduced bias in the outcome comparisons. When there are several factors being balanced simultaneously, these weights are typically derived from a regression model that predicts group membership based on that set of factors. The predicted probabilities from this model, sometimes called propensity scores (PS), are then used to derive inverse probability weights. The focus on methodological innovation in this area tends to be on how best to estimate the PS model. The intent is to move away from obtaining best model predictions and instead choose PS estimators that provide the best balance between the two groups being compared (Griffin et al., 2017). When successful, this balance ensures that the groups have the same distributions of covariates (in our case, matchable factors) when the PS weights are used in the analysis.

There is an abundance of weighting methods available to choose from (McCaffrey, Ridgeway, and Morral, 2004; van der Laan, Polley, and Hubbard, 2007; Westreich, Lessler, and Funk, 2010; Ridgeway et al., 2013; Imai and Ratkovic, 2014; Hainmueller, 2012; Zubizarreta, 2015). In general, the most promising methods continue to be those based on machine-learning methods (Lee, Lessler, and Stuart, 2010; Pirracchio, Petersen, and van der Laan, 2015). RAND commonly uses a robust machine-learning method called generalized boosted models (GBM), which aims to estimate a nonparametric model for the PS score that has much greater flexibility than a parametric model like the logistic model and which allows for substantially more covariates to be include than traditional parametric approaches. Machine-learning methods such as GBM have been shown to outperform logistic regression in terms of estimating high-quality propensity score weights (Westreich, Lessler, and Funk, 2010; Lee, Lessler, and Stuart, 2010).

In the end, we opted to use weighting because we wanted to retain the entire Army sample and include many factors in our model. We also wanted to use a method that would work with both simple matches (using just age and gender) as well as more complex matches (using all matchable factors) within each calendar year. Fundamentally, standardization, matching, and weighting are all mathematically equivalent methods if all matches are perfect and everyone has a match on the factors available. We note that both matching and weighting require having individual-level data to implement robustly, making them useful for this project. Although we did not use a matching approach, we use the term "matchable" factors to refer to the set of factors for which we are aiming to balance the groups and make the two groups comparable when using our PS weighting methodology.

Creating a Weighted NVDRS-CPS Sample

RAND Arroyo Center developed two sets of matching factors for this task. The first set of "standard factors" included commonly used demographic factors (i.e., age and gender) to weight the NVDRS-CPS sample to have comparable distribution to the Army population within each calendar year between 2003 and 2015. The second set of "augmented factors" included age and gender plus additional matchable factors identified in Chapter Two (age, gender, race/ethnicity, education, and marital status). We compared suicide rates between the Army and these two weighted NVDRS-CPS samples to assess the impact that including factors beyond the standard set have on the estimated NVDRS-CPS suicide rate and to assess the difference in suicide rates between the Army and NVDRS-CPS samples.

Weighting Method

Within each of the 13 years observed for this study, we used PS weights to weight the NVDRS-CPS sample to the Army sample. The target group of interest comprised active-component soldiers in a given year; the weighted comparison group consisted of individuals from the NVDRS-CPS sample in the same calendar year. To illustrate, the Army has far more men than the NVDRS-CPS sample (e.g., 85.9 percent versus 48.8 percent in 2015), and the Army sample in a given year is younger than the NVDRS-CPS sample (mean age 29.0 versus 44.6 in 2015). Since we want to estimate the suicide rate among the general population who are similar to members of the Army, we estimated PS weights for the NVDRS-CPS sample that "downweight" cases who look nothing like a soldier and upweight those that do, in each year. This type of weight is referred to as Average Treatment on the Treated (ATT) weights (Wooldridge, 2002).

We used the Toolkit for Weighting and Analysis of Nonequivalent Groups (TWANG) package in R to estimate PS weights that balance the two groups on each set of factors (standard and augmented), including interactions among the included factors. The TWANG package estimates PS weights for binary treatments/exposure indicators (such as Army versus NVDRS-CPS) using GBMs, a flexible, nonparametric estimation technique that can regress the group indicator (Army versus NVDRS-CPS) onto a large number of key matching factors (Ridgeway et al., 2017). GBM was first proposed for estimating PS weights by McCaffrey, Ridgeway, and Morral (2004) to circumvent the long and tedious process of fitting and refitting logistic regression models to obtain high-quality PS weights (Rosenbaum and Rubin, 1983; Dehejia and Wahba, 1999). Burgette, McCaffrey, and Griffin (2015) provide a brief overview of GBM, and we refer an interested reader to that document for details.

The ability of the PS to serve as a balancing score—that is, that results in groups with similar distributions on all observed covariates used in the PS model (Rosenbaum and Rubin, 1983), implies that, conditional on it, the distribution of the matching factors (age and gender for our standard factors; and age, gender, race/ethnicity, marital status, and education for our augmented factors) should be the same for the Army and NVDRS-CPS. While this is true in theory, in practice obtaining high-quality weights can be difficult. Thus, an important step in any PS weighted analysis is to assess the overall ability of the PS weights to create distributions of the covariates that are the same between the different groups being compared.

In the current study, comparability between the Army and NVDRS-CPS groups after PS weighting is assessed using standardized mean differences or effect sizes (ESs) (Ridgeway et al., 2017). For each covariate, the ES difference or standardized bias equals the absolute value of the difference between the mean for the Army and the PS weighted mean for the NVDRS-CPS sample divided by the standard deviation (SD) of the characteristics in the Army. More specifically, for covariate $k = 1, \ldots, K$,

$$ES_k = |\bar{x}_{k1} - \bar{x}_{k0}| / s_{k1},$$

where $\bar{x}_{k1}$ and s_{k1} are the unweighted mean and SD for the Army and $\bar{x}_{k0}$ is the ATT weighted mean for the NVDRS-CPS group. Thus, an ES tells us how large the differences are between the two groups on the mean of a given covariate. An ES of 0 implies no difference, while an ES equal to 1 implies a difference of 1 standard deviation. The ES is also commonly called Cohen's *d* (Cohen, 1992), and there are expected cut points for what constitutes a small (< 0.20), moderate (0.40), or large (> 0.60) difference between two groups. In our study, we aimed to have all absolute ES differences after applying weights to be below 0.10.

Weighting to 2015 Army Characteristics

Recent work from Watkins et al. (2018) recommended a different weighting strategy that weights both NVDRS-CPS and Army populations to the Army population in a single year. Thus, the Army suicide rates are also adjusted over years so that they reflect an estimate of the suicide rate for a given population of soldiers (e.g., the characteristics of soldiers in 2015). This eliminates variation over time in the suicide rate within the Army when that variation can be explained by shifts in the characteristics of soldiers over time. Watkins et al. refer to this as a "standard Army population" approach.

We performed this version of the analysis as well and show findings for the analysis in Appendix F. As noted, the two approaches do not result in different inferences about suicide risk in the Army compared with the NVDRS-CPS sample. This is because the population characteristics of both groups (e.g., the distribution of age and gender) have been relatively constant over the years included in our study. As such, we proceed by doing comparisons within year for the Army and NVDRS-CPS samples for the main part of this report.

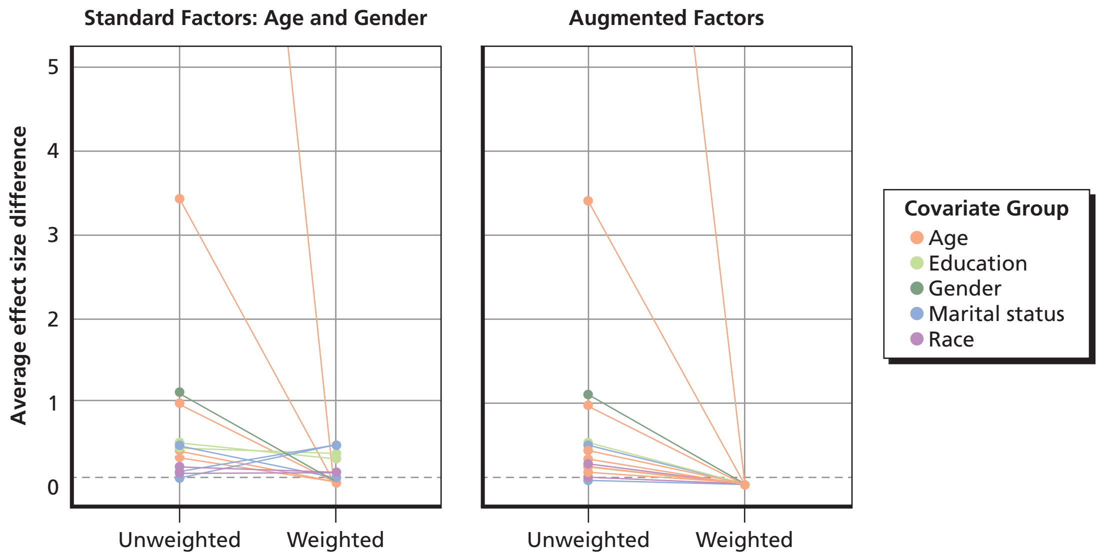
Comparing Suicide Rates

To compare suicide rates between the Army and NVDRS-CPS sample within each calendar year, we carefully examined three estimates of the "general population" rate: (a) the raw unweighted suicide rate in the NVDRS-CPS sample, (b) the weighted suicide rate for the NVDRS-CPS sample that was PS weighted to match the Army on the standard factors (age and gender), and (c) the weighted suicide rate for the NVDRS-CPS sample that was weighted to match the Army on the augmented factors (age, gender, race/ethnicity, education, and marital status). We examined both how those estimates differed from one another and how the comparison between Army and NVDRS-CPS suicide rates changes as a function of which estimates are used.

Results

Figure 5.1 shows the balance information for each set of matching factors. Each graph includes one line for each factor that we used to make the two populations comparable (e.g., each age group—age groups 18–21, 21–24, and so on—has its own line, and each race/ethnic subgroup

Figure 5.1
Mean Effect Size Difference Across Years 2003–2015, Before and After PS Weighted When Matching on Only Age and Gender (left side) and All Matchable Factors (right side)



has its own line). The lines are color coded into overarching categories for the key factors in our analysis (age, gender, race/ethnicity, education, and marital status). The graph on the left shows what happens when we create weights that make the Army and NVDRS-CPS samples comparable only on age and gender, while the graph on the right shows what happens when we create weights that make the two groups comparable on all factors.

As both graphics show, before weighting, the mean ES difference between the Army and NVDRS-CPS sample per year ranged from 0.06 to 16.35, suggesting very large differences between the two populations with respect to age, race/ethnicity, marital status, education, and gender. The largest differences occurred with age (ES > 3 for age groups over 55) and gender (ES = 1.08). After applying our PS weights using all matchable factors (age, gender, race/ethnicity, marital status, and education—the graphic on the right), these differences mostly disappeared, resulting in maximum mean ES of 0.47 when weighting on the "standard" factors (i.e., age and gender) and < 0.005 when using the augmented set. However, when we match the samples only on age and gender (as shown in the left-hand graphic), only age and gender are well-balanced, with mean ES = 0 across all years. In contrast, race/ethnicity, education, and marital status have lingering imbalances with mean ES well above our 0.10 threshold. Additionally, in cases such as marital status we actually see the ES go **up** when using weights that only account for age and gender.

Table 5.1 provides detailed information for Army and NVDRS-CPS samples in 2015, both before and after weighting for our two sets of matching factors. As shown, the two samples differ substantially on virtually all factors shown before weighting. The Army had a much smaller percentage of females than the NVDRS-CPS sample (14.1 percent versus 51.2 percent). The Army also has a smaller percentage of non-Hispanic whites than the NVDRS-CPS sample (58.7 percent versus 68.0 percent) and lower education levels on average (e.g., 77.1 percent of the Army had less than a bachelor's degree versus 58.1 percent of the NVDRS-CPS sample). The Army population is also younger than the NVDRS-CPS sample and has a larger

Table 5.1
Detailed Balance Information for Army Versus NVDRS-CPS in 2015

	Army	Unweighted NVDRS-CPS	Weighted NVDRS-CPS: All Matchable Factors	Weighted NVDRS-CPS: Age + Gender
Gender				
Female	14.1%	51.2%*	14.1%	14.1%
Race				
White	58.7%	68.0%*	58.7%	62.2%
Black	21.4%	12.8%*	21.5%	13.5%
Hispanic	12.9%	11.1%	12.9%	14.8%*
Asian/Pacific Islander	5.8%	5.8%	5.8%	6.6%
Other/unknown	1.2%	2.4%*	1.2%	2.9%*
Education level				
<BA	77.1%	58.1%*	77.1%	64.4%*
BA	15.3%	30.1%*	15.3%	28.0%*
>BA	7.7%	11.8%*	7.6%	7.6%
Marital status				
Married	60.1%	53.4%*	60.1%	33.3%*
Never married	34.2%	30.5%	34.2%	60.5%*
Formerly married	5.7%	16.1%*	5.7%	6.1%*
Age group				
Ages 18–21	12.7%	5.2%*	12.7%	12.7%
Ages 21–24	24.3%	7.9%*	24.3%	24.3%
Ages 25–29	22.5%	9.7%*	22.5%	22.5%
Ages 30–38	25.9%	16.3%*	25.9%	25.9%
Ages 39–44	9.7%	10.5%	9.6%	9.7%
Ages 45–54	4.6%	19.2%*	4.6%	4.6%
Ages 55–64	0.3%	18.2%*	0.3%	0.3%
Ages 65–75	0.0%	13.0%*	0.0%	0%

* Denotes when differences between the NVDRS-CPS sample have an absolute effect size difference > 0.10.

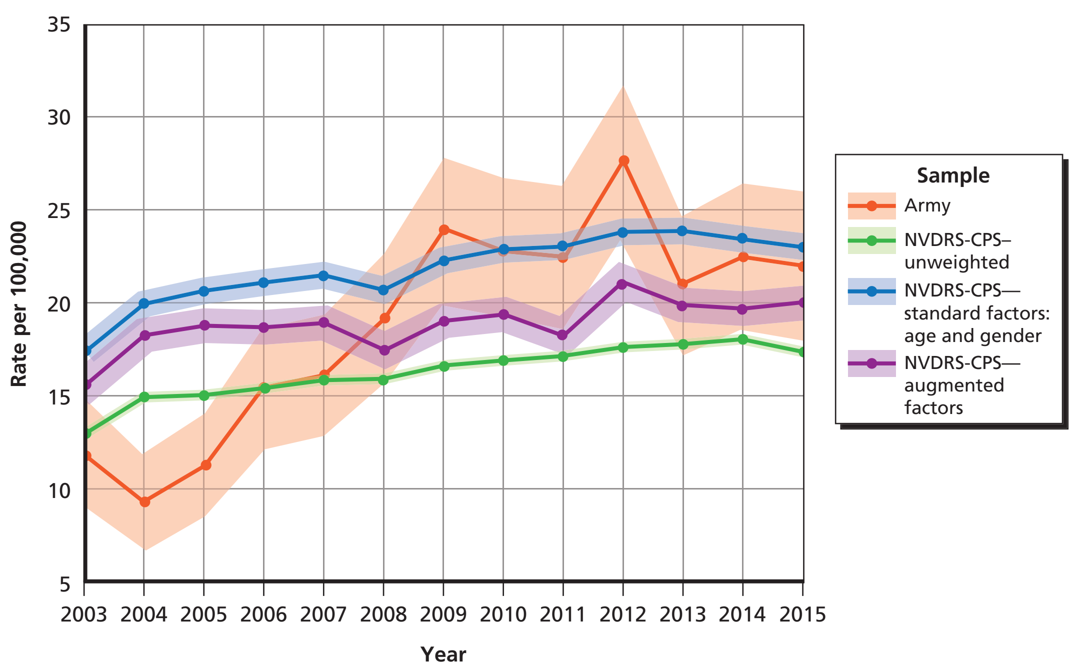
percentage of married couples (60.1 percent versus 53.4 percent) and fewer formerly married individuals (5.7 percent versus 16.1 percent) than the NVDRS-CPS sample. These trends were replicated across all years (detailed balance tables are available on request). After weighting for all augmented factors (age, gender, race/ethnicity, education, and marital status), the two populations have virtually identical distributions.

When weighting on just the standard factors of age and gender, the two populations look identical on age and gender but still have notable imbalances with respect to race/ethnicity, education, and marital status. Given the potential role race/ethnicity, education, and marital status play in explaining Army and NVDRS-CPS suicide rates, imbalance on these factors after matching on age and gender suggest that weights based on the augmented set of factors may be needed to properly identify the differences in suicide rates between the Army and NVDRS-CPS samples.

Figure 5.2 shows suicide rates over time for the Army as well as the unweighted NVDRS-CPS sample, the weighted NVDRS-CPS suicide rates using age and gender, and the fully weighted NVDRS-CPS suicide rate (using the augmented set of factors of age, gender, race/ethnicity, education, and marital status). The figure shows 95 percent confidence bands for each point estimate (suicide rate) based on using independent estimates of the variability within each year for each sample. As such, the bands represent conservative ways to eye for significant differences between suicide rates within each year. We note that comparisons over multiple years would be able to show significant differences even when the estimates in a single year are not different from one another.

As shown, weighting NVDRS-CPS to be similar to the Army on just the age and gender distributions increases the "expected" suicide rate in the weighted NVDRS-CPS sample. This likely occurs because the distribution of NVDRS-CPS has shifted to include a much higher percentage of males and a younger population, both of which are correlated with higher suicide rates. When comparing Army and NVDRS-CPS curves for this case, we note that before 2008, the Army suicide rate is significantly lower than the NVDRS-CPS suicide rate after adjusting for age and gender. After 2008, the confidence bands for these two curves (Army

Figure 5.2
Army and NVDRS-CPS Suicide Rates for Unweighted NVDRS-CPS, Traditional NVDRS-CPS Adjustment (Age Plus Gender), and Fully Adjusted NVDRS-CPS Samples



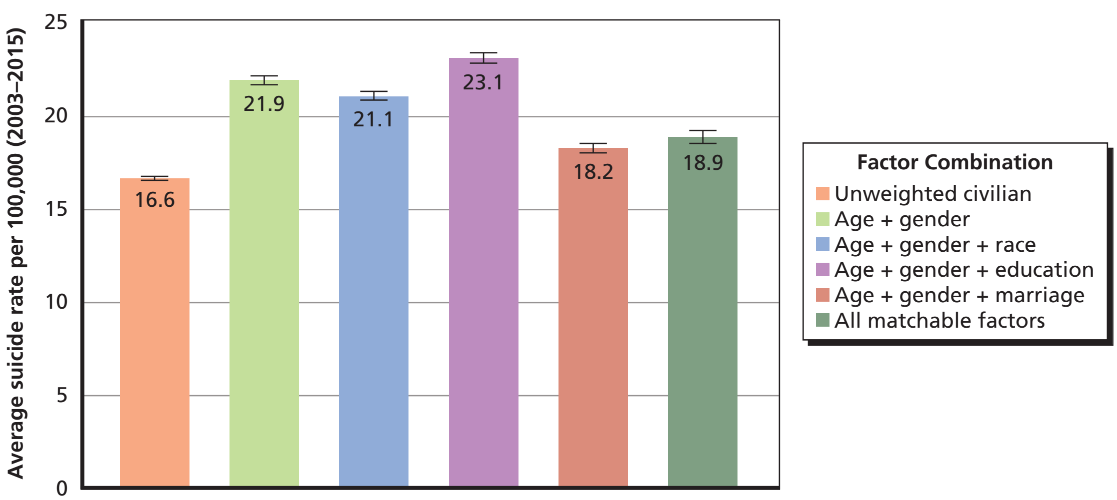
versus NVDRS-CPS adjusted for gender and age) are largely overlapping, suggesting little differences between the two populations. Additionally, the rates are shown to be close for the most part (except in 2012).

When we expand the weighting to include the augmented set of factors (race/ethnicity, education, and marital status), we notice that the "expected" suicide rate in the weighted NVDRS-CPS sample is now consistently reduced and lower in each calendar year than the NVDRS-CPS curve, which only used age and gender in the adjustment. In this fully adjusted analysis, the Army again has significantly lower rates in 2003, 2004, and 2005 than the weighted NVDRS-CPS sample. Confidence bands for the two curves generally overlap after 2005 (except in 2012), though the Army consistently has notably higher rates of suicide than the fully adjusted NVDRS-CPS sample. This suggests potential evidence of higher average rates in Army if one were to test across years rather than within years as shown in the graphic.

To gain a better understanding of which factor or factors (race/ethnicity, education, and marital status) appear to be driving the downward shift in NVDRS-CPS suicide rates after weighting with the augmented set of factors, we estimated three additional propensity score weights using slightly different sets of covariates that weighted the NVDRS-CPS sample to the Army, using age, gender, and (1) race/ethnicity, (2) education, and (3) marital status. These findings are shown in Figure 5.3. There is a very strong influence of marital status in lowering the weighted NVDRS-CPS suicide rate compared with weighting solely on age and gender; when matching only on age and gender within each calendar year, the NVDRS-CPS weighted suicide rate is 21.9 suicides per 100,000. When matching on age, gender, and marital status, the NVDRS-CPS weighted suicide rate decreases to 18.2 suicides per 100,000. On average, including the variable marriage decreases the expected suicide rate by 3.7 suicides each year compared with the model that only includes age and gender.

Recent work by Watkins et al. (2018) recommended that comparisons between the general U.S. and Army population suicide rates be standardized so that both the general U.S. and Army samples are all matched to one year of the Army. We repeated these analyses weight-

Figure 5.3
Average Suicide Rate per 100,000 in the Matched NVDRS-CPS Sample as a Function of Different Matching Factors



ing our NVDRS-CPS and Army samples to the 2015 Army. Thus, the Army suicide rates are also adjusted over years so that they reflect an estimate of the suicide rate for a given population of soldiers (namely, the characteristics of soldiers in 2015). This eliminates variation over time in the suicide rate within the Army when that variation can be explained by shifts in the characteristics of soldiers over time. These findings are shown in the Appendix F and provide relatively similar findings to those shown from the year-by-year analysis above. Adding the augmented factors in matching NVDRS-CPS to the Army produces lower suicide rates among the matched NVDRS-CPS relative to just matching on age and gender. The effect of marital status appears to be the strongest factor driving these findings.

Conclusions

As shown in Figure 5.2, the conclusions one can draw from comparing the Army and NVDRS-CPS suicide rates change depending on which factors are used to weight the NVDRS-CPS sample within each year. The preferred comparison should be viewed as the one that includes the factors that (a) are associated with suicide risk, (b) differ between military and the general population, and (c) are outside the control of the Army. These factors definitely include age, sex, and race/ethnicity. Additionally, it may also include education and marital status if these are not likely to be influenced directly by military service. Ideally, it would include all known and unknown factors as well. The sensitivity of the comparisons with the different sets of matching factors occurs because the set of factors used to weight the NVDRS-CPS sample impacts its underlying suicide rate. Here, we show how sensitive the comparisons can be to a trivial set of factors with the important caveat that even the comparisons shown are subject to bias because we cannot fully match the Army to the general U.S. population on all known and unknown risk factors for suicide.

Without any adjustment, the Army's suicide rate was generally lower than the national rate until 2006; in 2008 it surpassed the national rate. However, such differences are hard to interpret given the ways in which the Army differs from the NVDRS-CPS sample on multiple factors that are known to be correlated with suicide risk. Adjusting it so that the NVDRS-CPS sample is similar to the Army with respect to only age and gender (the most commonly done adjustment in the field) within each year, the NVDRS-CPS suicide rate increased to be almost comparable with the Army between 2009 and 2015. This occurs because the distribution of the NVDRS-CPS sample has shifted to include a much higher percentage of males and a younger population. However, this comparison is based on a weighted NVDRS-CPS sample that still differs from the Army with respect to race/ethnicity, education, and marital status, and thus is still suboptimal.

Fully adjusting for age, gender, race/ethnicity, educational attainment, and marital status within each year notably decreases the weighted NVDRS-CPS suicide rate from an analysis that uses only age and gender. When comparing the Army with the NVDRS-CPS sample that matches on the augmented set of factors covariates, the Army again has significantly lower rates in 2003, 2004, and 2005 than the weighted NVDRS-CPS sample. Confidence bands for the two curves generally overlap after 2005 (except in 2012), though the Army consistently has notably higher rates of suicide than the fully adjusted NVDRS-CPS sample, suggesting potential evidence of higher average rates in Army if one were to test across years rather than within years as shown in the graphic. More detailed analyses show that matching on marital status contributes the most to the decrease in estimated NVDRS-CPS suicide rates relative to the estimates that solely adjust for age and gender.

It is important to consider the potential impact that omitted variables might have on our comparisons in this chapter. As highlighted in Chapter Two, we identified several unmatchable factors from the literature review which we cannot use to match our Army and general populations. These included not only obvious things like geography, parenthood, occupation, mental health, and firearm availability but also things that are more intangible that predict both selection into Army service and suicide.

The main contribution of this report is on assessing available data sources for measuring observed suicide risk factors in the general U.S. population and the impact of standardizing the two populations using different sets of commonly available factors. We do not intend to estimate causal effects of serving in the Army. By understanding more fully how the comparisons shift depending on which set of factors are used to standardize the general population to look like the Army, we aim to show the impact of different sets of adjustment factors. It is inherent in all the provided estimates that unobserved factors are not matched on and therefore the comparisons in suicide rates being shown all still have important lingering bias that cannot ever be fully accounted for given real world data limitations on both populations

Adjustment via standardization—as we attempt to do in this study through comparison with a matched general population—aims to ensure that comparisons are not being explained by the matching factors themselves (e.g., such as age and gender differences in the population). Unfortunately, tackling such an analysis in a robust way is a complex task. The key problem in addressing the above questions is that the kinds of people who join the Army and the kinds of experiences to which Army life exposures a service member change over time and are related to each other. As noted in the introduction, the kinds of people who join the Army during times of economic downturn are different from the kinds of people who join the Army as a patriotic act in the wake of terrorist acts. Additionally, the experiences that are unique to Army service are different during times of war and times of peace. The research question would be much easier if the same selection mechanisms were at work during times of war and peace, as between-cohort differences in exposures to suicide experiences could be used to study the effects of time-varying experiences on time-varying within-service suicide rates. But that is not the case. An additional complexity exists in that we have strong selection out of service on the basis of risk factors for suicide with close to 20 percent of new enlistees leaving the Army in the first year of service, the majority of them for reasons that are related to risk of suicide, with the suicide rate among people with Army service being highest among people who left service prematurely in their first term of service.

It is highly certain that the difference between the Army and NVDRS-CPS suicide rate reported in this chapter will vary as a function of these unobserved/unmatchable factors, though it is unclear by how much. The analysis of unmatchable Army-relevant factors described in Chapter Three provides a partial assessment of the potential impact of unobserved covariates. However, as noted in Chapter Three, the substantive implications of how not being able to adjust for unmatchable factors impacts differences in Army-general population suicide rates will depend on the factor's effect size associated with suicide risk in each population as well as on the proportion of each population exhibiting the significant factors. It is very difficult to understand the size of these relationships for our unmatchable factors.

CHAPTER SIX

Conclusions

The Army has typically compared its suicide rate with the general population rate while adjusting for differences in the age and gender distribution of the two groups, as well as for temporal (yearly) differences. Our analysis confirmed that these are important factors to account for in any effort to compare the two groups. By examining different sources of general population data and applying a novel strategy of fusing two general population data sources (the NVDRS and CPS), we also identified three additional characteristics that differ substantially between the Army and general population and are related to suicide. These factors—race/ethnicity, educational attainment, and marital status—should also be considered for use in matched analyses of the two populations.

We wish to reiterate the overarching challenges in any comparison between the Army and the general U.S. population. Unfortunately, such a task is ultimately impossible given the number of unmeasured factors on which Army service members differ from the general U.S. population. Even if one were able to assemble a set of covariates that could be balanced in a way that explains the difference in the suicide rates in the Army and civilian sectors of society, that does not mean that we have shown that the kinds of people who join the Army are not different from those who do not join with respect suicide risk factors or that Army experiences do not increase risk of suicide. We know that unmeasured covariates exist, making it difficult to fully understand how suicide risk and rates compare between the Army and the general U.S. population.

Based on our analysis, we make the following recommendations.

1. **Given that comparisons will be made between the Army's suicide rate and that of the general population, these comparisons should adjust for age, gender, and year, and for the additional matchable factors of race/ethnicity, educational attainment, and marital status.**

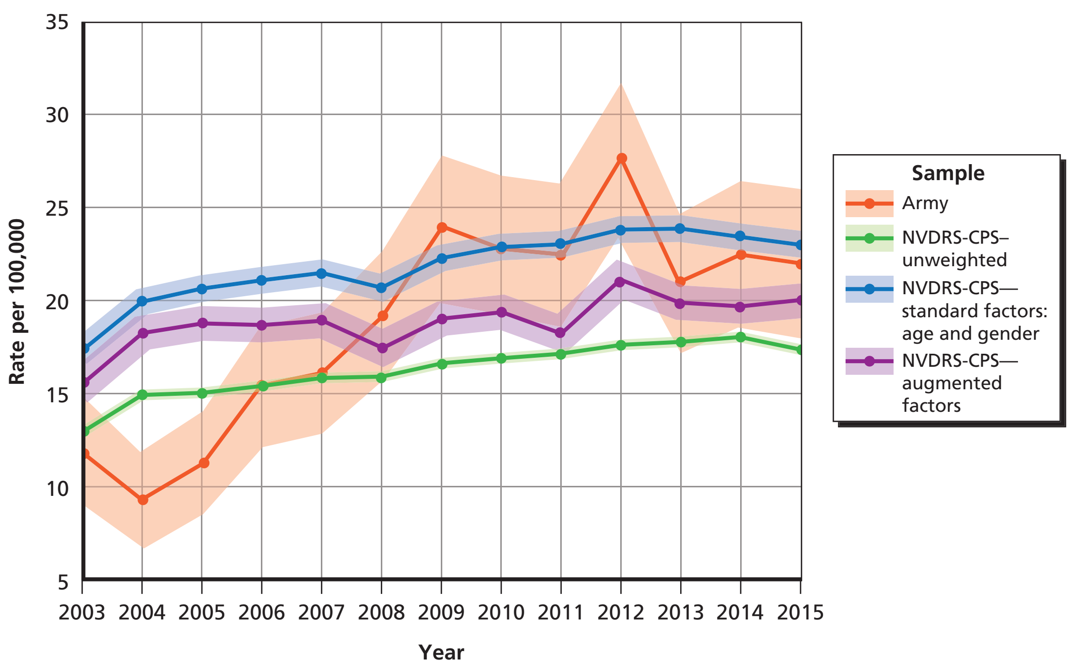
As shown in Figure 6.1 (reproduced from Figure 5.2), before *any* adjustment, the Army's suicide rate was generally lower than the NVDRS-CPS sample rate until 2006. In 2008 it surpassed the rate in our NVDRS-CPS sample. Adjusting the NVDRS-CPS sample to be similar to the Army based solely on age and gender distributions *increases* the suicide rate in the NVDRS-CPS sample because the distribution of the NVDRS-CPS sample has shifted to include a much higher percentage of males and a younger population. The Army's suicide rate was generally lower than the adjusted NVDRS-CPS rate (i.e., the "expected rate") until 2009, at which point the two rates were *generally comparable* and remained so until 2015.

Further adjusting for race/ethnicity, educational attainment, and marital status *decreased* the adjusted NVDRS-CPS rate. This changes the comparison with the Army such that in 2008 the Army's rate *exceeded* the fully adjusted suicide rate and *remained higher* until 2015. Additional analyses suggest that matching the NVDRS-CPS sample to the Army on marital status plays the largest role in shifting the suicide rates for the NVDRS-CPS weighted sample after matching on age and gender.

As noted, accounting for factors such as race/ethnicity, educational attainment, and marital status notably shifted the estimated suicide rate for the NVDRS-CPS population in large part due to the fact that we have changed the underlying Army population characteristics to which we are trying match the general population, affecting the conclusions one might draw from the Army-civilian comparison. The decision to include marital status and education depends primarily on the underlying assumed theory about the relationship between Army service and these factors. If the theory for both is that the Army attracts individuals based on their education levels and marital status/aspirations, then both should be included in the matched comparisons between the Army and the general population. If education and/or marital status is being driven to change by DoD policies, it would be best not to include them directly in the matching because their inclusion might obscure the impact of serving in the Army on suicide risk.

The Army differs from the general population in many ways, and it is these differences that account for the shift in the expected suicide rates described above. More specifically, we identified five additional factors (geography, parenthood, occupation, mental health, and firearm availability) that could be important for comparing and studying the two groups. However, limited data availability constrained our ability to include these unmatchable factors in

Figure 6.1
Army and Civilian Suicide Rates for Unweighted NVDRS-CPS, Traditional NVDRS-CPS Adjustment (Age Plus Gender), and Fully Adjusted NVDRS-CPS Samples



our analysis. Given the known geographic variation in suicide rates across states, it may be fruitful to include state-level indicators as more states participate in the NVDRS and those data become available. However, there are theoretical reasons why both where the soldier died and where the soldier was from may be important to consider.

We also explored Army-specific factors of rank and deployment history that are associated with suicide even after controlling for our matchable factors. Our analyses of unmatched factors showed that unit location (CONUS or not in CONUS), rank, deployment history, and soldier's job explain variation in suicide risk above and beyond the matchable factors considered in our analysis.

Job-related duties and operational tempo are other factors that may distinguish soldiers from the general population. Drawing parallels between general population and Army job categories would have been a challenging task in and of itself. However, we were unable to even consider this challenge due to limitations in general population mortality data systems. If the Army is interested in pursuing adjustment by occupation or industry, it should consider the following recommendation:

2. **The Army should collaborate with the U.S. Census Bureau, the Centers for Disease Control and Prevention, and the Department of Labor to improve occupation/industry coding for general population deaths.**

Given that occupation is a known risk factor for suicide in both populations, better quality data on the general U.S. population would be useful to obtain. In Appendix A, we describe our efforts to better understand how occupation and industry are coded in federal data systems. Because the CPS data on industry and occupation are coded by the U.S. Census Bureau, we recommend applying the Census Bureau's coding procedures, which involve both autocoding via an algorithm and manual coding by clerks for difficult entries. RAND Arroyo Center made several attempts to collaborate with Census Bureau researchers to apply their autocoder to the NVDRS data we had acquired. We were not successful due to the proprietary nature of their coding algorithm and the Census Bureau's limited resources. A collaboration between the Army, Census Bureau, CDC, and Department of Labor could make facilitating more accurate coding of death records a priority.

A soldier's job-related duties and operational tempo ("unmatchable" factors) are other factors that may distinguish the Army and general populations. However, we were unable to draw parallels between general population and Army job categories due to limitations in how occupation is coded in the mortality data available on the general population. A collaboration between the Army, Census Bureau, CDC, and Department of Labor could increase the priority assigned to more accurate coding of occupation and industry in death records for the general population. Additionally, extensive work would be needed to determine how to determine which general population occupations best align with military occupations. Preliminary work in this area has been done by Wenger and colleagues (2017) and could be used as a basis for this work.

3. **The Army should collect voluntary data on soldiers who own personal firearms and should encourage the CDC or another federal agency to resume collecting voluntarily provided survey data on gun ownership and use in the general population.**

This recommendation echoes recommendations that RAND has made more generally to improve gun policy science (Morral, 2018a). Soldiers may differ from their general population counterparts regarding ownership of or access to personally owned firearms, the suicide method used in the majority of Army suicides. Adjusting for this factor may also be important for making comparisons between the Army and general population. As noted, this factor falls into our third set of characteristics for which careful consideration before inclusion in the standardization process. If the goal of an analysis is to estimate the effect of Army service on suicide risk, including such controls as access to firearms should be avoided because this characteristic represents the specific mechanism by which Army service affects suicide risk (e.g., through availability of military firearms). However, there may be some research purposes for which one does want to control for firearm access—for example, to directly study how much of the effect of Army service on suicide risk is mediated through access to military firearms. Unfortunately, high-quality data in both the general population and the Army is fundamentally lacking. The lack of data on personally owned firearms among soldiers and the general populations impedes the Army's ability to adjust for or study a potentially important factor that may distinguish soldiers from members of the general population and that is correlated with suicide.

Finally, the Army and the general population may differ with respect to the presence of mental health conditions that are among the strongest risk factors for suicide. Unfortunately, limited data in the general population on either diagnosed mental health conditions or mental health symptoms make weighting on this factor impossible in the study's data sets. However, the Army could examine the rate of suicide among those with mental health diagnoses relative to individuals from the general population with the same diagnoses, adjusting for the sociodemographic characteristics described above. Our final recommendation, therefore, is that

4. **Future research should examine suicide risk among those with mental health diagnoses in the Army relative to similar individuals in the general U.S. population.**

The Army and general U.S. population may differ with respect to mental health conditions, which are among the strongest risk factors for suicide. For example, the 2014 STARRS report showed that lifetime prevalence estimates of a variety of mental disorders were significantly higher among new soldiers who were surveyed during their first few days after reporting for duty than similarly matched individuals from the U.S. population on age, gender, education, and race/ethnicity in 2011–2012. This highlights the underlying challenge in comparing the Army's suicide rate with the general U.S. population in that the kinds of people who joined the Army are generally at substantially higher risk of suicide related to history of mental disorders than similar individuals from the general U.S. population who could have, but did not, enlist. Even if these types of differences do not extrapolate to all years (e.g., it might also be that all those new soldiers with high burden of prior psychopathology never made it past their first year of service and did not contribute to the high Army suicide rate), there is a great need to be able to match the two populations on mental health to be able to better understand the role mental health plays in suicide rates and the comparison of rates between the Army and the general U.S. population.

Data deriving from medical claims may be most easily linked to death data and has detailed information on mental health diagnoses and thus may be the most fruitful avenue for future research. The Army could replicate the methods used in this study to examine the rate of suicide among those with mental health diagnoses in the Army relative to individuals in the general population with the same diagnoses, adjusting for the sociodemographic characteristics described above. Such research will likely require partnership with an existing health system or data system like the National Inpatient System that not only reports mental health diagnoses within an insured population but also links mental health information to cause of death data. To do this, data on diagnoses will be needed not just for suicide cases but also for the entire Army and general U.S. populations at risk. Additionally, the Mental Health Research Network is a consortium of 11 health care systems that serves over 11 million individuals across 11 states. Its data have been linked to cause of death data (Ahmedani et al., 2014) and may be suitable for conducting such an analysis. This investigation may provide as much insight into the phenomena of Army suicides and how they compare with general population suicides as it does about the type of care provided in the Military Health System, a recent priority for the Department of Defense (Hepner et al., 2018) and the Army (Srinivasan et al., 2016). Again, to do this, data on diagnoses will be needed not just for suicide cases but also for the entire Army and general populations at risk. Additionally, care will need to be taken to address any differences in general population and Army/military health systems regarding coding of psychiatric diagnoses.

Conclusion

Comparing the Army's suicide rate with that of the general population is important to help determine whether suicides in the Army are related to Army experiences or a function of the people the Army enlists (i.e., young men with high school degrees who are likely to marry before their peers). The comparison can yield insights into Army suicides that may be critical for the Army Resilience Directorate. However, such comparisons do not change the Army's obligation to care for its soldiers, including preventing soldiers from killing themselves. Understanding how the Army compares with the general population community may help to identify strategies that can prevent suicide, but such analyses do not diminish the Army's responsibility to those who serve.

APPENDIX A

Industry and Occupation Coding in the NVDRS

In death records, industry and occupation are typically coded in "free-text" fields in which medical examiners or coroners write in their descriptions of the person's industry and occupation. Guidelines for doing so exist (Robinson et al., 2012). Since 2012, there have been nearly 40,000 suicide deaths in the United States; thus, the challenge for surveillance is to determine an efficient way to convert these text fields into standardized, numeric codes for surveillance. In addition, it is important that the codes used for deaths are comparable with the codes used in the CPS (which constitute our denominator) so that we can accurately estimate suicide risk. We describe below the options we pursued to efficiently code the NVDRS occupation and industry data.

Existing Coding Schema

There are four coding schema used to classify occupations and industries in U.S. data systems. These are described in Table A.1.

Table A.1
Coding Schema for Industry and Occupation

	Description	Use
North American Industry Classification System (NAICS), 2012	2- to 6-digit codes interpretable at multiple levels: sector (2-digit), subsector (3), industry group (4), NAICS industry (5), and National industry (6). Versions: 2002, 2007, 2012, 2017. Designed under the Office of Management and Budget to replace the now-defunct SIC (Standard Industrial Classification) 4-digit system.	Developed to provide an international standard to describe business establishments in statistical data publications from federal government agencies (in U.S., Canada, and Mexico).
Standard Occupation Code (SOC), 2010	2- to 6-digit codes, listing 23 major groups, 97 minor groups, 461 broad occupations.	Used by federal statistical agencies to classify workers into occupational categories.
Census Industry Codes	Based on 2012 NAICS, with reduced detail to avoid potential disclosure issues. 2-digit versions list 20 sectors; 4-digit versions list 269 industries	Used by Census Bureau to classify workers into industry categories, based on census/ACS responses. Can be crosswalked to 2012 NAICS.
Census Occupation Codes	Based on 2010 SOC. 2- through 4-digit codes. 2-digit versions list 23 major groups; 4-digit versions list 539 codes.	Used by Census to classify workers into occupational categories, based on Census/ACS responses. Can be crosswalked to 6-digit 2010 SOC.

NVDRS Occupation and Industry Codes

Some observations in NVDRS arrive with occupation and industry precoded by the states when they are submitted to the CDC, whereas other states have not coded occupation and industry at all. Even among states that have provided coding, their rates of comprehensiveness vary. Only Alaska and Utah have provided codes for almost all their records; South Carolina and Georgia often have missing information for the write-in text fields (Figure A.1).

Availability of Free-Text Job/Industry Data

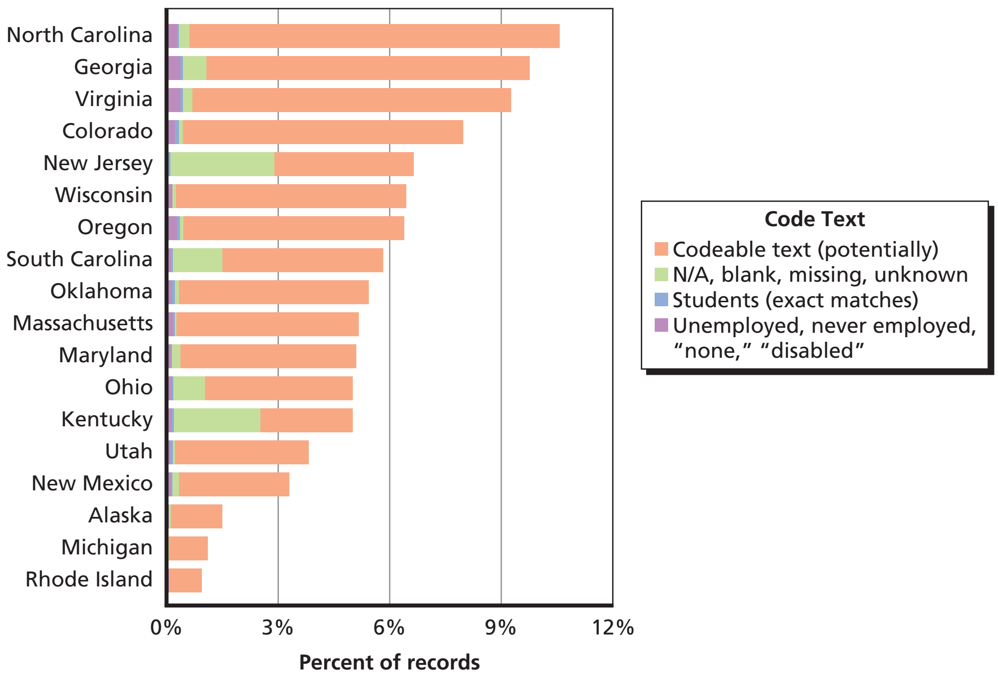
We examined the job/industry text fields in the NVDRS and grouped results into four categories: (1) unemployed/never employed/none/disabled; (2) students; (3) not available (N/A)/blank/missing/unknown; and (4) potentially codeable text (all remaining entries). As shown in Table A.2, in total 87.0 percent of industry and 83.5 percent of occupational fields fall into category 4 and thus eligible for coding.

Figure A.1
Coding of Occupation and Industry Codes in the NVDRS, by State

Table A.2
Disposition of Industry and Occupation Fields in the NVDRS

Code Text	Industry	Occupation
Codeable text (potentially)	87%	83.5%
N/A, blank, missing, unknown	9.6%	7.1%
Students (exact matches)	0.8%	5.7%
Unemployed, never employed, "none," "disabled"	2.6%	3.7%

Figure A.2
Disposition of Industry and Occupation Fields in the NVDRS, by State



While in general 87 percent of industry and 83.5 percent of occupation codes are categorized as "codeable," the quality of the free text varies by state, shown in Figure A.2. Note that in this figure, the x-axis refers to the percent of *all* records in NVDRS (so the entire chart sums to 100 percent and the length of the bar for each state represents what portion of NVDRS records it accounts for). As shown, states like South Carolina, Ohio, New Jersey, Kentucky, and Georgia have high proportions of fields that are N/A, blank, missing, or unknown.

Converting Text Fields to Industry and Occupation Codes

We identified four automated systems to convert text fields into one or more of the aforementioned industry/occupation coding schemes.

NIOSH Industry and Occupation Computerized Coding System

The National Institute for Occupational Safety and Health's Industry and Occupation Computerized Coding System (NIOCCS) is a web-based software tool developed by the CDC's National Institute for Occupational Safety and Health (NIOSH) to translate free-text information describing industry and occupation to standardized codes—that is, NAICS, SOC, and Census Bureau industry and occupation (I&O) codes. It takes two text fields as inputs: "job title" and "industry title." To control the rate of false positives, NIOCCS can be run with two different accuracy thresholds: the "high" accuracy setting is calibrated to provide 90 percent confidence in autocode results and the "medium" settings, 70 percent confidence.

As NIOCCS does not generate codes for all text entries provided, it is best used as part of a multiphased strategy. Performance of the NIOCCS with the 2012 NVDRS data by RAND Arroyo Center for this project is described below.

SOCcer

SOCcer was developed by Russ et al. (2016) to assign SOC-2010 codes based on free-text job description components and previously coded industry classifications. To generate codes, SOCcer takes a stacked ensemble approach, in which the submodels are used to generate predictions which are then input into a logistic regression that determines the final classification. Models were fit using a training data set of free-text job description data and corresponding manually coded SOC codes. The submodels used in the stacked ensemble approach include classifiers based on job title, industry, and task. Job title-based classifiers involve computing a soft Jaccard index which computes the relative frequency with which a free-text entry (after correcting for spelling irregularities) belongs to a given SOC code. An industry-based classifier is included to nudge occupation codes toward occupations that occur with more than 1 percent prevalence within any industry; for example, so the job title text "fireman" is assigned the "firefighter" occupation code (common in the "fire protection" industry) rather than a "kiln operator" code (a niche occupation within the pottery production industry). A task-based classifier uses a "fuzzy fingerprint" approach to associate words appearing in job task description with SOC codes. Last, predictions are input into a logistic regression that generates probabilities for various SOC codes, outputting as the autogenerated code that with the highest estimated probability.

Running SOCcer requires inputting two text strings describing "job title" and "task." With only these data inputs, Russ et al. (2016) demonstrate that SOCcer can effectively generate autocodes for a large number of cases. When comparing SOCcer against NIOCCS to reproduce expert assignments from the Occupational Safety and Health Administration's Integrated Management Information System, Russ et al. (2016) found that "the highest scoring SOC code from SOCcer using both job title and industry matched the expert consensus SOC code for 36.8 percent and 70.4 percent of the jobs at the 6-digit and 2-digit levels, respectively." In contrast, NIOCCS provided assignments for only 3,332 (28 percent) U.S. Renal jobs with the remaining assignments not provided because they had a "medium" level confidence or less. Overall, the agreement between NIOCCS and coders was only 17.3 percent at the 6-digit level (Russ et al., 2016) when considering nonassigned jobs as a mismatch.

U.S. Census Bureau Autocoder

Since 2012, the U.S. Census Bureau has used an autocoding algorithm to help code free-text descriptions of (a) employer name, (b) industry descriptions, (c) job description, and (d) job duties.[1] This approach is two-phased: autocoding is attempted first, with difficult entries referred to manual coding by clerks.

To calibrate the threshold for whether an entry should be autocoded or referred to manual coding, a study comparing outcomes of manual and autocoding was performed. As such, the algorithm classified 56 percent of industry codes and 43 percent of occupation codes.

[1] Employer name (INW2): "For whom did this person work?" Industry Description (INW3): "What kind of business or industry was this? Describe the activity at the location where employed." Occupation description (OCW1): "What kind of work was this person doing?" Job duties (OCW2): "What were this person's most important activities or duties?"

Manual/autocoder comparisons found clerical disagreement rates of 4.5 percent and 5.9 percent, respectively. Ten percent of industry codes (representing 1.3 percent of records) are not even attempted to be autocoded, and it's the same with 20 percent of occupation codes (1.6 percent of records).

The census method follows the following steps: (1) assemble data dictionaries from common strings of consecutive words, (2) fit a logistic regression prediction I&O codes from presence of wordbits (total counts and conditional probabilities), (3) use "hard-coding," overriding the logistic regression output, for text entries that commonly generate errors from the logistic regression. More detail about these steps is described as supplemental text at the end of this appendix.

RAND Arroyo Center made several attempts to collaborate with Census Bureau researchers to apply their autocoder to our NVDRS data. However, the Census Bureau declined to share either (a) the algorithm for their logistic regression or (b) the dictionaries they use to hard-code certain entries. RAND Arroyo Center also attempted to coordinate a strategy whereby we would send them our NVDRS data for them to auto- and/or manually code. They were willing to consider this in principle, but in practice they seemed much too overburdened to take on additional work.

Performance of Various Autocoding Algorithms

We applied both the NIOCCS and SOCcer algorithms to the 2012 NVDRS occupation and industry fields, both of which are available via an online tool (see CDC, undated c; National Institutes of Health, undated). Specifically, we applied both the NIOCCS high and medium threshold algorithms (the latter classifies more uncertain entries) as well as the SOCcer algorithms with a score 50 percent cutoff and without a score cutoff (classifying all entries, regardless of uncertainty). We generated, where possible, industry/occupation codes of several types: NAICS code, Census Industry Code, Census Occupation Codes, SOC (6-digit), and SOC (2-digit). The NIOCCS algorithm directly outputs Census Bureau I&O Codes, NAICS codes, and SOC codes. SOCcer only provided estimates for 2-digit SOC codes; in a personal communication, a developer of SOCcer told us: "[T]he ability to code to 6-digit SOC-2010 is a function of the quality of the provided job information. In many studies we have found insufficient job detail with which to code (based on expert review) beyond a 3-digit level." The results are summarized in Table A.3.

These results demonstrate the difficulty of using autocoding algorithms in isolation to generate industry/occupation codes for observations in the NVDRS data set. There is a sharp trade-off between prediction accuracy and comprehensiveness. Under recommended accuracy settings, NIOCCS (under "high" accuracy) generated predictions for only 45 percent of NVDRS observations and the SOCcer (calibrated for "50 percent" accuracy) for only 19.5 percent.

A full evaluation of these algorithms requires a better understanding of the accuracy-comprehensiveness trade-off. The availability of manually coded entries in some states provides the opportunity to validate algorithms based on their performance for these labeled entries. We compute rates of agreement with the manually coded 2-digit SOC entries, comparing them with the SOCcer-generated codes (using the 50 percent threshold versus no threshold) and the NIOCCS-generated codes (at the "medium" and "high" calibrations). Though one might

Table A.3
Percent Data Availability by Code

Code Type	NIOCCS High	NIOCCS Medium	Manual	Soccer (50% Threshold)	Soccer (No Threshold)
U.S. Census industry	67.9	83.6	9.0	N/A	N/A
U.S. Census occupation	67.6	83.5	9.1	N/A	N/A
NAICS industry	46.4	62.2	N/A	N/A	N/A
SOC (6-digit)	45.2	61.1	N/A	N/A	N/A
SOC (2-digit)	45.2	61.1	7.9	19.5	1

expect to find an arc-shaped receiver operating characteristic (ROC) curve trading off accuracy and comprehensiveness, instead we find that accuracy matched against manually coded observations is quite poor and inelastic with respect to calibration. This suggests that either the manual codes are often incorrect or that the nature of the free-text entries in the NVDRS data set are not amenable to the autocoding algorithms.

To investigate further, we implement within-state analysis of the rates of agreement between manually coded entries and autocoded observations (Figure A.4.). We see that agreement rates vary wildly by state: for Alaska, Colorado, and New Mexico, the 2-digit SOC codes generated by NIOCCS (medium accuracy) and SOCcer (50 percent threshold) match the manually assigned codes with between 25 percent and 38 percent accuracy. Five states (Kentucky, Massachusetts, Wisconsin, Utah, and North Carolina) yielded accuracies between 5 percent and 10 percent. For all other states, occupation codes are either not provided to NVDRS or are provided but are probably low quality, as they never matched the NIOCCS/SOCcer-generated codes. Meanwhile, a comparison of 2-digit SOC codes output by NIOCCS

Figure A.3
Manual Code-Agreement Versus Availability by Algorithm, Calibration (2-Digit SOC)

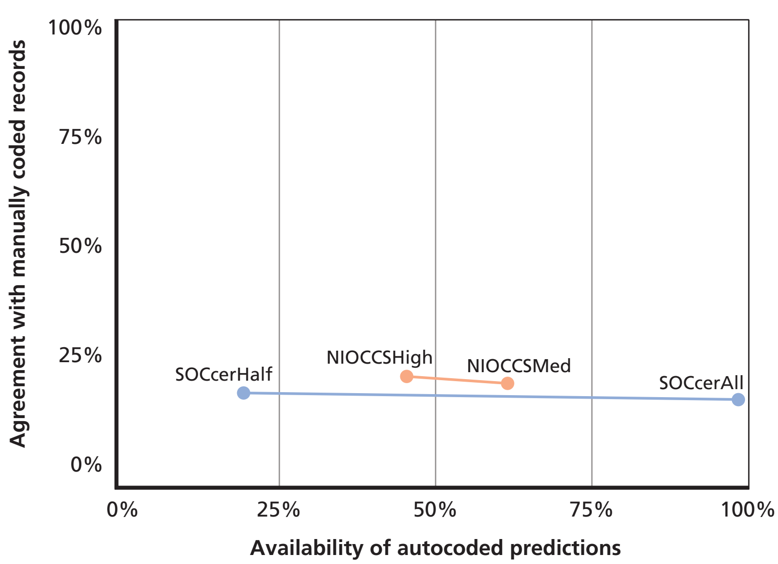
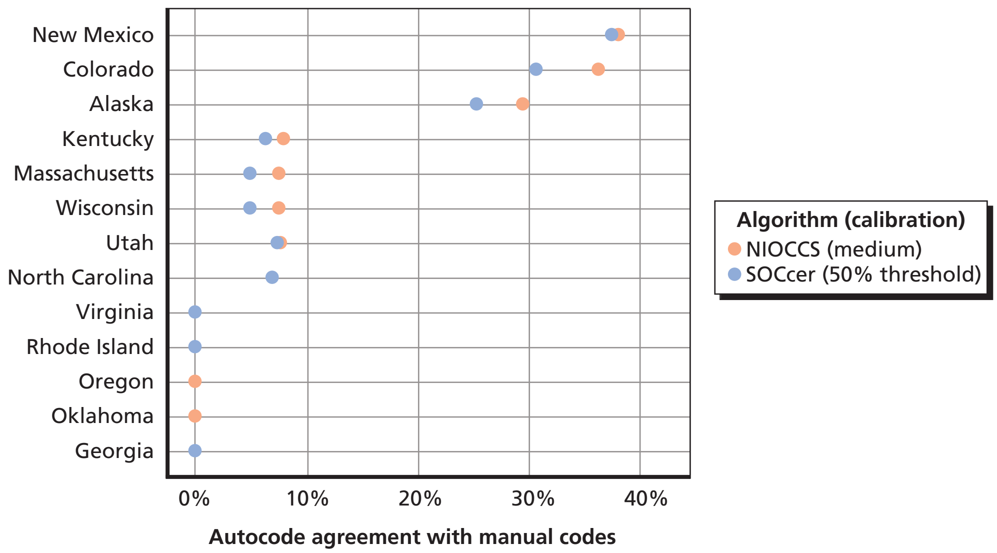


Figure A.4
Two-Digit SOC Autocode Agreement by State, Method



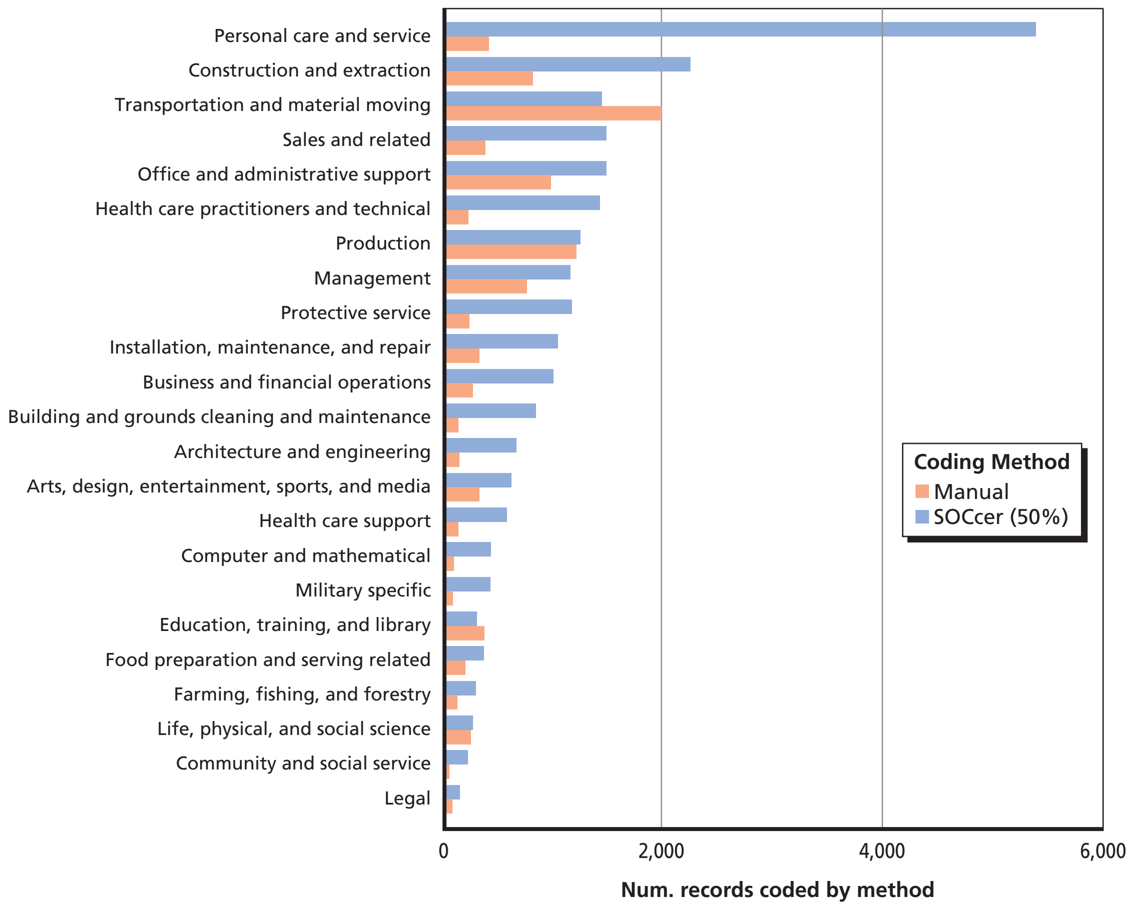
and SOCcer shows an average intra-algorithm agreement rate of about 80 percent, which holds robustly across different states. A combination of these two factors (the high rate of agreement between NIOCCS and SOCcer codes, plus the large state-to-state variation in agreement across manual codes and autogenerated codes) suggests that there are substantial data quality concerns with either the manually coded SOC codes or the accompanying free-text descriptions.

Given the potential data quality concerns, an attempt to describe the distribution of occupations in the NVDRS data would benefit by reference to both manual codes provided by the states and autogenerated codes. Figure A.5. illustrates the distribution of occupation codes by each of these methods. Note the different patterns: SOCcer identifies the two most popular occupation groups to be "personal care and service" and "construction and extraction"; the state-provided codes instead point to "transportation and material moving" followed by "production."

Conclusion

Efforts to code industry and occupation based on free-text description of job title, task, and industry title are under development by a number of actors, including U.S. federal agencies and leaders in the academic community. Substantial progress has been made in the past decade, but nonetheless autocoding methods are best used as combination of a multiphased approach, coupled with semiautomatic methods (e.g., conversion of text strings to codes using a "hard-

Figure A.5
Distribution of Available 2-Digit SOC Codes by Manual, SOCcer Methods



coded" dictionary matching strings to codes) and manual clerical coding. Application of two leading algorithms, NIOCCS and SOCcer, to the NVDRS data set suggest that this data set is particularly difficult for the production of autocoded entries. Both NIOCCS and SOCcer algorithms are generally poor at providing autocoded entries, and where those entries are able to be produced, they do not validate well with the SOC codes that some states provide to the NVDRS system. The combination of these factors suggests potential data concerns with the SOC codes provided by states and potentially also the write-in text describing job and industry title. In conclusion, analysis of the occupation and industry codes in the NVDRS data set should include nonautomated methods and perhaps also include careful validation of state-provided codes.

Supplementary Detail on the Census Autocoder

Data dictionaries are constructed that match "wordbits" to I&O codes, working from a sample of 1.5 million records out of the 2.3 million that were manually coded in the 2010 census. Wordbits are consecutive word strings (not cleaned for spelling errors or prefixes and suffixes).

Separate dictionaries are constructed for one-word, two-word, and all-word length word bits. Separate dictionaries are also constructed to match each of the four text fields (INW2, INW3, OCW1, OCW2) to I&O codes. For consideration in inclusion in a data dictionary, word bits must (a) occur at least 30 times in the data, and (b) be associated with a single I&O code on at least half of its appearances. Likewise, "crosscode" dictionaries are created that associate INW2 and INW3 text with occupation codes or OCW1 and OCW2 text with industry codes. Eligibility requirements for crosscode are different: appearing 30 times in the data but matching a single code on at least three-quarters of occasions. For prediction purposes, crosscodes receive less weight than same-code matches.

Use of logistic regression. Industry codes are assigned first, though with consideration of crosscodes from occupation fields; afterward, occupation codes are assigned with consideration of the autoassigned industry code. This mimics the clerical process, where coding decisions are made together. Explanatory variables include information from data dictionaries (e.g., wordbit count [i.e., sample size for the word bit in the data set], code frequency percent [percent of a word bit's appearances that return a given I&O code]) and ACS response data (e.g., sex, age, number of words in INW2 entry). Except for the frequency, percent continuous variables are binned (information on how is not provided, but bins were calibrated to maximize agreement rates). Logistic regressions are used, predicting one if the autocoded entry matches the clerically coded entry, outputting a prediction score (P-HAT).

Use of hard codes. Text entries that most commonly generated errors received follow-up attention by experts, who created "hard codes" that map text strings to I&O codes. (Hard-coded entries, or exact word string matches, are treated as if they are certainly accurate, i.e., receiving a regression output P-HAT of one.)

Use of thresholds for issuing an autocode. Thresholds were calculated separately for industry and occupation codes, with the goal of only issuing autocodes just as good as clerical coding (i.e., matching a more careful "expert coding" at the same rate as a standard clerical code, for about 5-percent agreement). Autocodes were sorted by P-HAT, generating a ROC curve, at which separate P-HAT cutoffs were assigned for I&O codes. The cutoffs were set at levels that, on the 2010 data, generated codes for 43 percent of occupation records and 56 percent of industry records.

Problems with autocoding performance: words with multiple meanings: for example, manager/dealer/editor; excessive weight on some words: for example, x-ray installer; phrasing: "engineering manager" versus "engineer, managing a team"; spelling.

Potential improvements. Input could be done with a drop-down menu rather than free text. Data could be cleaned to correct spelling errors, remove prefixes/suffixes, or combine synonyms into single entries.

APPENDIX B

Suicide Modeling Methods

Cumulative Logistic Models

Cumulative (person-level) logistic regression models assess the relationship between a set of risk factors and a specified dichotomous outcome event. The logistic regression model assumes that the length of follow-up time has limited effect on parameter estimates, which may not be valid in mortality studies given that the likelihood of observing a death will tend to increase with the length of time an individual is followed or with the overall length of time that the individual is observed over the study period. Prior simulation studies have shown that the cumulative logistic model can produce serious bias when length of follow-up varies considerably across subjects (Annesi, Moreau, and Lellouch, 1989; Ingram and Kleinman, 1989).

Estimates from these models express the association between risk factor and suicide as relative odds. Given our interest is in the risk of suicide as opposed to the time to event, logistic models are on their face most appropriate to our outcome of interest. However, standard logistic regression analyses may be inappropriate when individuals are studied for varying lengths of time as models need to take into account not only whether the event of interest occurred for an individual but also the length of time that each individual is exposed to the risk of the event.

Weighted Logistic Models

A limited literature has used a weighted logistic regression to account for differential follow-up length. We identified one study using a weighted logistic model in estimation of the recurrence of adenomas, where the weight function was an increasing function of follow-up length (Hsu, Green, and He, 2007). The authors evaluated three weighting functions: (1) a uniform weight function (= observed follow-up divided by maximum follow-up); (2) an exponential weight function; and (3) a nonparametric weight function derived from a Kaplan-Meier survival curve).

Using Monte Carlo simulations, their article showed that traditional cumulative logit methods underestimated event rates across all scenarios and that the Kaplan-Meier survival method (with right-endpoint imputation) tended to overestimate event rates. The performance of the Kaplan-Meier method (with midpoint imputation) depended on the hazard function, and the weighted logit method (even with the simpler uniform weighting function) provided reasonable estimates for the event rate.

Survival Analysis Models

Survival analysis methods may be more appropriate in the study of Army suicide risk due to two primary reasons. First, these methods are more appropriate in the presence of censored observations (i.e., allows analysis before all events have been observed). Second, these methods can accommodate the staggered entry of individuals into the data set (i.e., individuals enter the data set at different times during the study). Note that a key assumption is that right-censoring is noninformative—that is, event times are independent of the censoring mechanism (analogous to the "missing at random" assumption).

Additionally, survival analysis methods are preferred over logistic models when there is interest in the time to event as opposed to just the occurrence of the event. However, this is not the focus of our current study and is thus a less important consideration when selecting between models. Indeed, unlike the cumulative logistic model, survival models have as the outcome of interest "time to event," and it has been noted that the effect of a particular covariate on the cause-specific hazard function of a particular failure type may differ greatly from the effect of that covariate on the corresponding cumulative incidence function (see Zhang et al., 2009).

Discrete Time Survival Analysis (DTSA)

The person-time logistic model (Ingram and Kleinman, 1989), partial logistic regression (Efron, 1988), or pooled logistic regression (D'Agostino et al., 1990) are methods of DTSA that modify the cumulative logistic regression model by expressing the dependent variable as the number of outcome events per person-time unit rather than per person. It appears these may all be equivalent models just with different naming conventions. With the person-month as the unit of analysis, individuals would thus only contribute information for the months in which they are observed in the data set, thus allowing for analyses with censored data or staggered entry. Advantages of the approach are that (1) time-varying predictors are easily accommodated, (2) random effects can be included, (3) competing risks can be examined, and (4) non-proportional hazard models can be estimated (Koslow et al., 2014). Time-dependent effects are introduced as interactions between the covariates and the discrete factor (or set of dummy variables) representing time. Gibbons et al. (2003) apply these methods to study transplant and mortality risk using a framework that includes a vector of random effects to model variations across Organ Procurement Organizations.

In the person-time logistic regression, estimated coefficients can be interpreted as discrete time survival coefficients (Willett and Singer, 1993). The assumption of the person-time logistic model is that the probability of suicide death for an individual in any time interval is independent of the number of time intervals already survived—that is, survival time is assumed to be exponentially distributed (Anderson et al., 1980; Ingram and Kleinman, 1989). Efron (1988) shows that the person-time logistic model approximates standard continuous parametric models of the survival hazard. In further simulation studies, this model has been shown to provide a good approximation to the Cox time-dependent regression model when the rate of occurrence of the event of interest is low, the follow-up period is short, and the relative risks associated with risk factors included in the models are small or moderate (see Callas et al., 1998).

Poisson Models

Like the person-time logistic model, the Poisson regression model assumes exponential survival time. One potential advantage of using the Poisson regression over the person-time logistic is that it provides maximum likelihood estimates of the rate ratio (or hazard ratio) that are consistent with traditional methods for analysis of cohort data (Ingram et al., 1990). In grouped data for use in Poisson regressions, an offset variable can be included to account for total exposure. However, Poisson regression analysis can also be conducted on ungrouped person-time data where each observation represents a unit of person-time at risk (Loomis et al., 2005). A recent example of the use of Poisson regression to analyze mortality with follow-up time used as an offset is offered in Cheng et al. (2016). These are types of modified Poisson regression models with robust and/or sandwich variance estimators for the estimation of risk ratios (see Zou, 2004; Zou and Donner, 2011). Previous studies have noted that the Poisson regression model is equivalent to the piecewise exponential survival model where time (or duration of exposure) is partitioned into separate intervals and the baseline hazard is assumed constant within each interval (Holford, 1980; Laird and Olivier, 1981). Thus, one advantage of discrete time survival models compared with the piecewise exponential survival model or Poisson model is that one does not need to make the assumption that the hazard function is constant within each interval.

Link Functions in DTSA

Discrete time survival analysis using the logistic link and person-month as the unit of analysis appears to be the preferred model used in recent studies of suicide risk factors with Army STARRS (Schoenbaum et al., 2014; Gilman et al., 2014; Kessler et al., 2015b). However, there are other potential link functions for use in the generalized linear model, such as the probit link function and complementary log-log link function (Rodriguez, 2008). An advantage to the latter is that the resultant regression coefficients are equivalent to those of an underlying proportional hazards regression model (Austin, 2017). Thus, the estimated coefficients can be interpreted as having a relative effect on the hazard of the occurrence of the event. Additionally, the complementary log-log link has been proposed as being more appropriate when time is continuous but only observed in grouped discrete form. Rodriguez (2007) notes the following recommendations regarding the choice of survival analysis methods using generalized linear models.

- If time is truly discrete (e.g., an event that can only occur at discrete points of time, such as an outcome of grade completion which only occurs at the end of a school year), then one should probably use the discrete model with a logit link, which has a direct interpretation in terms of conditional odds and is easily implemented using standard software for logistic regression.
- If time is continuous but one only observes it in grouped or interval form, then the complementary log-log link would seem more appropriate. Results based on the complementary log-log link should be more robust to the choice of categories than results based on the logit link.
- If time is continuous and one is willing to assume that the hazard is constant in each interval, then the piecewise exponential approach based on the Poisson likelihood is preferable. This approach is reasonably robust to the choice of categories and is unique in allowing the use of information from cases that have partial exposure.

- If time is truly continuous and one wishes to estimate the effects of the covariates without making any assumptions about the baseline hazard, then Cox's partial likelihood is a very attractive approach.

Split Population Models (or Cure Models)

In standard survival analysis methods, an implicit assumption is that the cumulative distribution function approaches one as time at risk becomes sufficiently large—that is, all individuals who do not experience suicide death during the study period would eventually die of suicide if observed over a long enough period (Schmidt and Witte, 1989). This assumption may be inappropriate in the context of suicide risk as some individuals may never commit suicide, regardless of external circumstances or length of follow-up. To allow for the existence of a never-suicide population, there is a modified hazard approach that has been commonly used in the study of substance use initiation and cessation (see Kostova et al. [2016] for some citations): the split population duration (SPD) model, which allows a fraction of the right-censored observations to correspond to subjects who will never experience the event of interest. (Note: In the biostatistics literature, this may be referred to as the Mixture Parametric Cure Model [Lambert et al., 2007].) This model was originally applied (at least in the economics literature) by Schmidt and Witte (1989) to study risk of criminal recidivism. Swaim and Podgursky (1994) generalized the model to allow for right-censoring; Forster and Jones (2001) and Tsodikov et al. (1998) developed SPD models that accommodate time-varying covariates.

The split population model first estimates the individual probability of ever experiencing the event of interest, then weights the hazard function by this probability. Intuitively, split population models allow for a subpopulation that never experiences (and will never experience) the outcome of interest. The general approach is the estimation of a mixture of a standard hazard density and a point mass at zero; thus, compared with standard DTSA, split population models estimate an additional parameter(s) for the probability of eventual failure. Further, split population survival analyses allow the association of covariates with experiencing the outcome and time to experiencing the outcome to be examined separately within the same model.

See Lai and Yau (2009) or Seppä et al. (2010) for discussion of multilevel mixture cure models with random effects. Seppä et al. (2012) also provide potentially relevant discussion of quantifying realistic error margins for random error in numbers and proportions of avoidable deaths. See Xi et al. (2012) for an application of the Cox cure model to suicide risk.

Machine-Learning Techniques

While it is outside the scope of this study to review recent literature on machine-learning techniques in the estimation of suicide risk, we here point the interested reader to several relevant studies.

- *An Evaluation of Randomized Machine-Learning Methods for Redundant Data: Predicting Short and Medium-Term Suicide Risk from Administrative Records and Risk Assessments* (Nguyen et al., 2016)

- “Evaluating the High Risk Groups for Suicide: A Comparison of Logistic Regression, Support Vector Machine, Decision Tree and Artificial Neural Network” (Amini et al., 2016)
- “Support Vector Machine Versus Logistic Regression Modeling for Prediction of Hospital Mortality in Critically Ill Patients with Haematological Malignancies” (Verplancke et al., 2008)
- “Predicting Risk of Suicide Attempts over Time Through Machine Learning” (Walsh et al., 2017)
- “Developing a Practical Suicide Risk Prediction Model for Targeting High-Risk Patients in the Veterans Health Administration” (Kessler et al., 2017).

Candidate Data Sources on General Population Suicides

Table C.1
Candidate Data Sources on General Population Suicides

Data Set	Geographic Coverage	Access Requirements	Description
NVDRS-Public	6 to 27 states, depending on year	No access requirements	Descriptive data including demographic information and details around the death where available; 2003–2015; only accessible through online portal
NVDRS-Restricted Access Database (RAD)	6 to 18 states, depending on year	Proposal Package	Case-level file of the NVDRS database with all available variables provided for research and approved use
National Death Index (NDI)	Entire U.S.	Proposal Package & Payment	Data through 2015 now available; compressed mortality file—county-level national mortality file and county-level population file; county-level mortality file contains subset of variables available in the detailed annual mortality file; must request recent years
National Vital Statistics System (NVSS)	Entire U.S.	No access requirements	Provide mortality data by multiple cause of death for all deaths occurring within the United States. Each record in the microdata is based on information abstracted from death certificates filed in vital statistics offices of each state and District of Columbia
Census of Fatal Occupational Injuries (CFOI)	Entire U.S.; includes workplace suicides only	No access requirements	Pre-aggregated (e.g., industry by event or exposure, occupation by event or exposure, etc.)
Mortality Detail Files	Entire U.S.	No access requirements	Every death or fetal death registered per year in the U.S. from 1968–1992
Detailed Mortality (WONDER)	Entire U.S.	See CDC (undated a)	Provides aggregated summaries from CDC databases, including NVSS
Current Population Survey (CPS)	Entire U.S.	No access requirements	Monthly panel survey data weighted to be nationally representative; provides detailed information on general population characteristics

APPENDIX D

Data Harmonization

Common demographic variables such as race and education are often coded differently across separate data sources. In order to use disparate data sources together, the separate coding schemes must be harmonized so that all variables contain identical value sets. In this study, there were two main data harmonization efforts to ensure successful fusion and comparison across the CPS, NVDRS, and Army data. The first concerned aligning variable values for all matchable factors between the NVDRS and CPS prior to data fusion to create our NVDRS-CPS sample. The second harmonization effort was between the resulting NVDRS-CPS sample and the Army sample prior to propensity score weighting and analysis. In both instances, the goal of recoding was to leave variables as granular as possible across both data sets, which generally meant collapsing one data set's values of a variable up to match the other data set. Table D.1 presents the original variable values for the NVDRS and CPS as well as the harmonized value. Table D.2 presents the original variable values for the NVDRS-CPS sample and the Army sample along with the final, harmonized value.

Table D.1
Data Harmonization for Merging NVDRS and CPS to Impute Suicide Cases

NVDRS Race	Harmonized	CPS Race
Race		
White	White	White
Black	Black	Black/Negro
Asian/Pacific Islander	Asian/Pacific Islander	Asian only
		Hawaiian/Pacific Islander
		Asian-Hawaiian/Pacific Islander
American Indian	American Indian	American Indian/Aleut/Eskimo
Two or more races	Two or more races or other	White-American Indian
		Black-American Indian
		White-Black-American Indian
		White-Asian
		White-Black
		White-Hawaiian/Pacific Islander
		Black-Asian
		White-Black-American Indian-Asian
Other		White-American Indian-Asian
		White-Black-Asian
		White-Black-Hawaiian/Pacific Islander
		American Indian-Hawaiian/Pacific Islander
		White-American Indian-Hawaiian/Pacific Islander
		Black-American Indian-Asian
		Two or three races, unspecified
		Four or five races, unspecified
Unknown	N/A	N/A
Ethnicity		
Not Hispanic	Not Hispanic	Not Hispanic
Hispanic	Hispanic	Mexican
		Puerto Rican
		Cuban
		Dominican
		Salvadoran
		Other Hispanic
		Do not know
Unknown	N/A	Not available/no response

Table D.1—Continued

NVDRS Education Level	Harmonized	CPS Education Level
Education		
<= 8th grade	<= 8th grade	None or preschool Grades 1, 2, 3, or 4 Grades 5 or 6 Grades 7 or 8
9th–12th grade	9th–12th grade	Grade 9 Grade 10 Grade 11 12th grade, no diploma
High school or GED grad	High school diploma or equivalent	High school diploma or equivalent
Some college credit	Some college credit	Some college but no degree
Bachelor's	Bachelor's	Bachelor's degree
Associate's	Associate's	Associate's degree, occupational/vocational program Associate's degree, academic program
Master's	Master's	Master's degree Professional school degree
Doctorate	Doctorate	Doctorate degree
Unknown	N/A	NIU/blank
Marital Status		
Married/civil union/ domestic partner	Married	Married, spouse present Married, spouse absent
Never married	Never married/single	Never married/single
Single, not otherwise specified		
Widowed	Widowed	Widowed
Divorced	Divorced or separated	Divorced
Married, but separated		Separated
Unknown	N/A	N/A

NOTE: NIU = not in universe.

Table D.2
Data Harmonization for Merging Army and NVDRS-CPS Samples for Propensity Score Weighting

Army	Harmonized	NVDRS-CPS
Race		
Hispanic	Hispanic	Hispanic
White	Non-Hispanic white	White
N/A		N/A
Black	Non-Hispanic black	Black
Asian/Pacific Islander	Non-Hispanic Asian/ Pacific Islander	Asian/Pacific Islander
Other/unknown	Other/unknown	Two or more races or other
Education		
Less than bachelor's	Less than bachelor's	<= 8th grade 9th–12th grade High school diploma or equivalent Some college credit
N/A		N/A
Bachelor's	Bachelor's	Bachelor's
More than bachelor's	More than bachelor's	Associate's Master's Doctorate
Marital Status		
Married	Married	Married N/A
Never married	Never married	Never married/single
Formerly married	Formerly married	Widowed Divorced or separated

APPENDIX E

Analyses for Location and Deployment History

Our results in Chapter Five are limited to including variables that are available for matching in both the Army and the general population databases. As shown in Chapters Two and Three, there are likely to be other factors not in one or both databases that may also be important for matching. RAND Arroyo Center implemented a series of sensitivity analyses that explored the role other such factors might play in explaining differences (or similarities) observed between Army and matched general populations on suicide outcomes. More specifically, for unit location and deployment history, we implemented additionally stratified analyses that examine how the comparison of suicide rates change as a function of where soldiers are located (United States or non–United States) and deployment history. The results below show how the comparisons between the Army and matched general population might changes depending on these factors.

Location

Given the protective effect we found for soldiers located outside the United States, we implemented an additional analysis that matched the NVDRS-CPS sample to the Army after excluding soldiers overseas. Figure E.1 shows how the suicide rates for the Army compare when including versus excluding service members who are located outside the United States. As shown, exclusion of these soldiers has little impact on the observed suicide rate, suggesting that a comparison between NVDRS-CPS and the entire Army only would change slightly if these soldiers were excluded from comparisons.

Deployment History

Given that our analyses showed that soldiers with prior deployment history have higher rates of suicides on average than soldiers without any deployments, we implemented two stratified analyses that matched the NVDRS-CPS sample to the Army based on deployment history. Prior research has shown that the relationship between suicide and deployment history is complex, and as such, we used this analysis to explore whether there were time-varying effects of deployment on the comparison of suicide rates between the Army and NVDRS-CPS sample. As shown in Figure E.2, there are clearly complex relationships in this regard. Most notably, for all three curves, we first note that the confidence bands overlap, suggesting the inferences for comparing suicide rates between NVDRS-CPS and the Army do not differ greatly between the different groups. In terms of point estimates for the odds ratio comparing suicide risk between the Army and the NVDRS-CPS sample, we generally expected the green curve rep-

Figure E.1
Army Suicide Rates With and Without Soldiers Serving Overseas

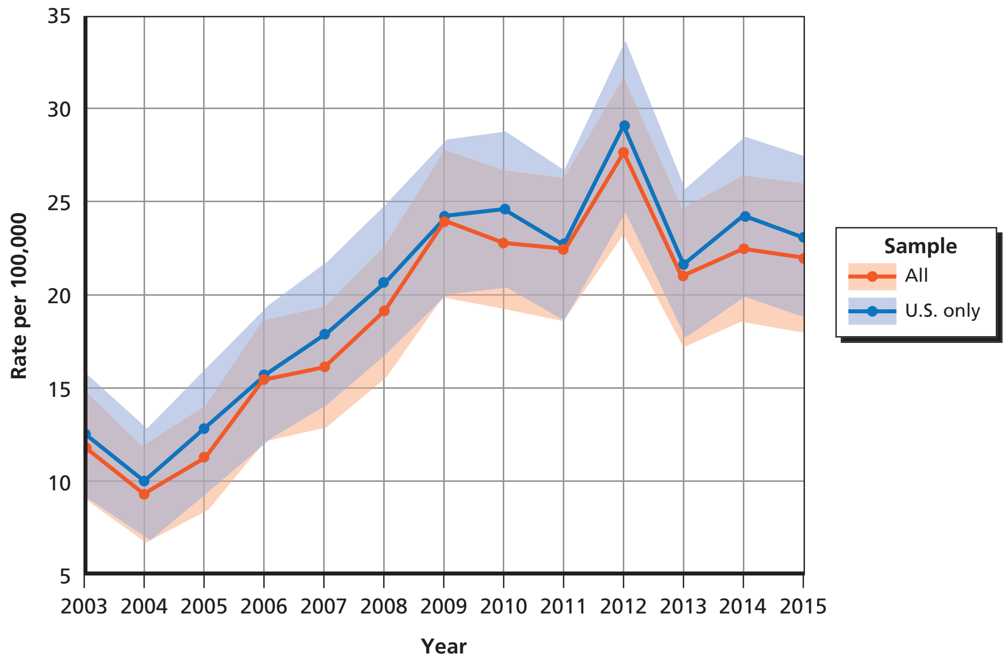
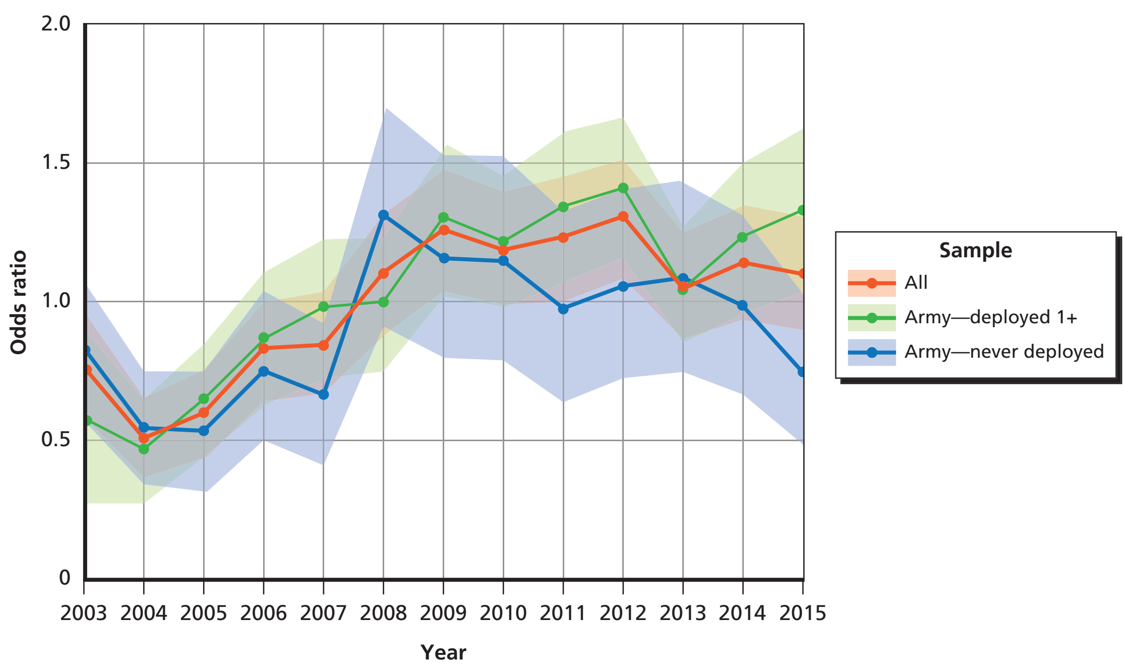


Figure E.2
Estimated Odds Ratio when Comparing Army to NVDRS-CPS Suicide as a Function of Deployment History



resenting the subset of the Army with a deployment history to be slightly higher than the red curve which plots the odds ratios for the overall population; this generally occurs. Conversely, we expected the odds ratios for the nondeployed soldiers to generally be lower than the overall population odds ratio; again, this is the observed pattern we expect in a few years.

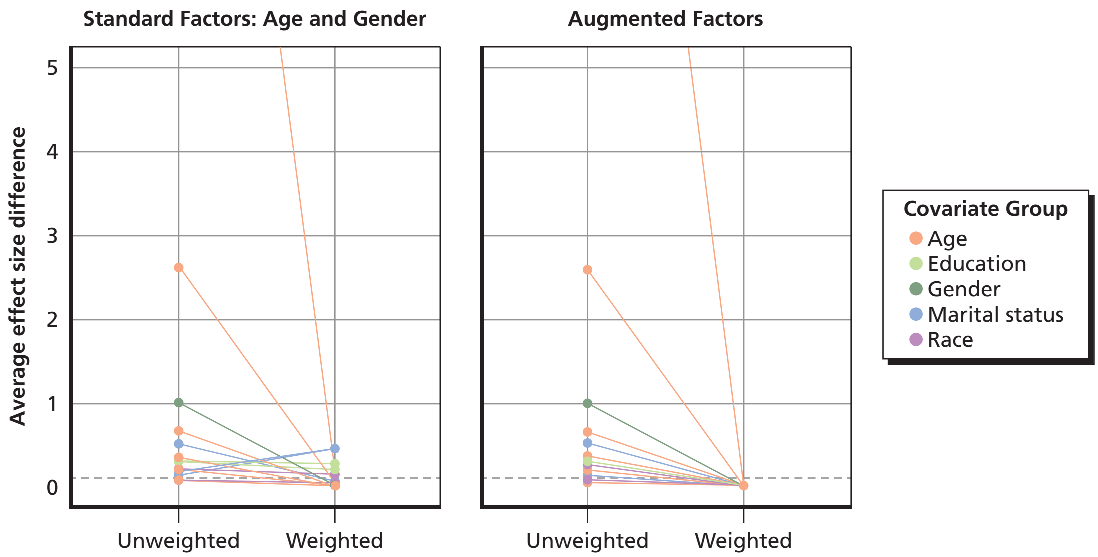
APPENDIX F
2015 Army Analysis

In this appendix, we present our findings from sensitivity analyses which aimed to weight both the civilian sample and the Army sample between 2003 and 2015 to one target population, namely the Army 2015 population. We undertook this set of analyses given recent work from Watkins et al. (2018) that strongly recommended that comparisons between the general population and Army population suicide rates be standardized so that both the general population and Army samples are matched to one particular year of the Army with 2015 designated as the current best target (Watkins et al., 2018). They refer to this as a "standard Army population" approach. As noted in the report, overall implications do not differ between using the standard Army population versus our primary analysis which weighted the NVDRS-CPS sample within each year to the Army population from that same year. As such, we present the analysis from the within-year analysis as our primary set of analyses given its ability to better reflect time trends. Here, we show in detail the findings from our standard Army population analysis.

Figures F.1 and F.2 show the balance information across years for each set of matching factors. In this case, two sets of graphics are displayed—one for weighting the NVDRS-CPS sample in each year to the 2015 Army target sample and the other for weighting the Army 2003–2014 samples to look representative of the 2015 Army. Each individual graph includes one line for each factor that we aimed to make the two populations comparable on (e.g., age groups 18–21, 21–24, and so on each have their own line and each race/ethnic subgroup has its own line). Then, the lines are color coded into overarching categories for the key factors in our analysis (age, gender, race/ethnicity, education, and marital status). The graph on the left in a given figure shows what happens when we create weights that only make the samples comparable on age and gender, while the graph on the right shows what happens when we create weights that make the groups comparable on all factors.

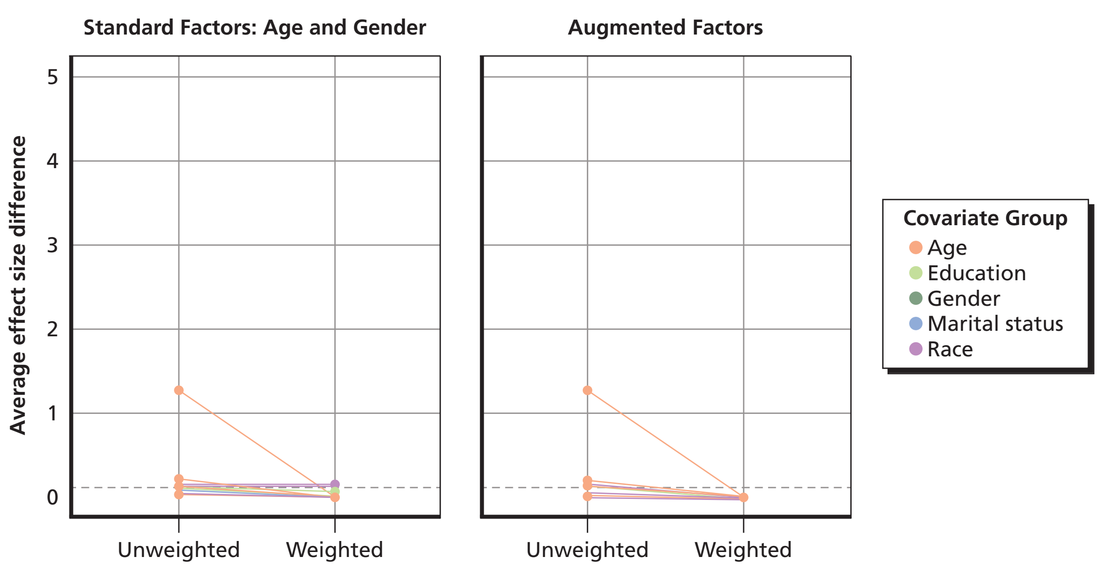
As shown in Figure F.1 and as is consistent with our findings in the main report, prior to weighting, the mean ES difference between the 2015 Army and NVDRS-CPS sample per year ranged from 0.06 to 14.89, suggesting clear, strong imbalances or differences between the two populations with respect to age, race/ethnicity, marital status, education, and gender. The largest differences occurred with age (mean ES > 2.6 for age groups over 55) and gender (e.g., mean ES = 1.00). After applying our PS weights which use all matchable factors to the NVDRS-CPS samples from each year, these differences greatly disappeared, resulting in a mean ES of 0.02 or 0.46 when matching on all matchable factors or just age and gender, respectively. When we match the samples on only age and gender, both age and gender are well balanced with mean ES equal to zero across all years but with race/ethnicity, education, and marital status having lingering imbalances with mean ES well above our 0.10 threshold.

Figure F.1
Mean Effect Size Difference Between the NVDRS-CPS Samples Across Years 2003–2015 Versus the 2015 Army, Before and After PS Weighted When Matching on Only Age and Gender (left side) and All Matchable Factors (right side)



In contrast, Figure F.2 shows the balance information across years for each set of matching factors when matching our Army samples from 2003 to 2014 to the Army in 2015. Here, we can see that there were much smaller differences between the 2015 Army and the Army in most years. The mean effect size difference ranges from is 0.03 to 1.28; the greatest differ-

Figure F.2
Mean Effect Size Difference Between the Army Samples Across Years 2003–2014 Versus the 2015 Army, Before and After PS Weighted When Matching on Only Age and Gender (left side) and All Matchable Factors (right side)



ences occurring with age (mean ES = 1.28 for the age group 65–74, for which there are generally very few to no soldiers in a given year). All other covariates have mean ES less than 0.17 prior to weighting. The generally small-to-moderate effect sizes differences are not surprising given that demographics of the Army have changed over time (as can be seen in Table 3.2). The weighting that used all matchable factors removed these differences so that each weighted Army sample looks representative of the 2015 Army in our analysis. Again, after applying our PS weights to the Army samples from 2003 to 2014, the small-to-moderate differences disappeared, resulting in maximum mean ES of 0.00 or 0.17 when matching on all matchable factors or just age and gender, respectively. Even in the Army analysis, when we match the samples on only age and gender, both age and gender are well balanced with mean ES equal to zero. In contrast, race/ethnicity, education, and marital status have lingering imbalances with mean ES well above our 0.10 threshold.

Table F.1 provides detailed balance information for the Army and NVDRS-CPS samples from 2009, both before and after weighting to the 2015 Army. As shown, there are larger differences between the NVDRS-CPS 2009 sample and the 2015 Army sample than between the two Army samples. Similar to the results show in Table 4.4, the 2015 Army, prior to weighting, had a much smaller percentage of females than the 2009 NVDRS-CPS sample (14.1 percent versus 51.7 percent). The 2015 Army also has a smaller percentage of non-Hispanic whites than the 2009 NVDRS-CPS sample (58.7 percent versus 71.5 percent) and lower education levels (e.g., 77.1 percent of the Army has less than a bachelor's degree versus 62.3 percent of the 2009 NVDRS-CPS sample). The 2015 Army population is also meaningfully younger than the 2009 NVDRS-CPS sample and has a larger percentage of married people (60.1 percent versus 55.7 percent) and less formerly married (5.7 percent versus 18.5 percent) than the 2009 NVDRS-CPS sample. When comparing the 2009 to 2015 Army samples, the 2009 Army sample had a smaller percentage of non-Hispanic whites than the 2015 Army sample (51.5 percent versus 58.7 percent), a higher percentage of soldiers with less than a bachelor's degree (82.8 percent versus 77.1 percent), and a higher percentage of Asian/Pacific Islander and other race solders. These types of trends were replicated across all years (detailed balance tables available on request).

After weighting for all matchable factors (age, gender, race/ethnicity, education, and marital status), the 2009 NVDRS-CPS and Army samples have virtually identical distributions to the 2015 Army on all factors. When weighting on just age and gender, the two samples continue to look identical to the 2015 Army on age and gender but have several lingering imbalances with respect to race/ethnicity, education, and marital status for the 2009 NVDRS-CPS sample and with respect to race/ethnicity and education for the Army 2009 sample. Given the potential role race/ethnicity, education, and marital status play in explaining Army and NVDRS-CPS suicide rates, these types of lingering imbalances for our model which matches on only age and gender should be concerning.

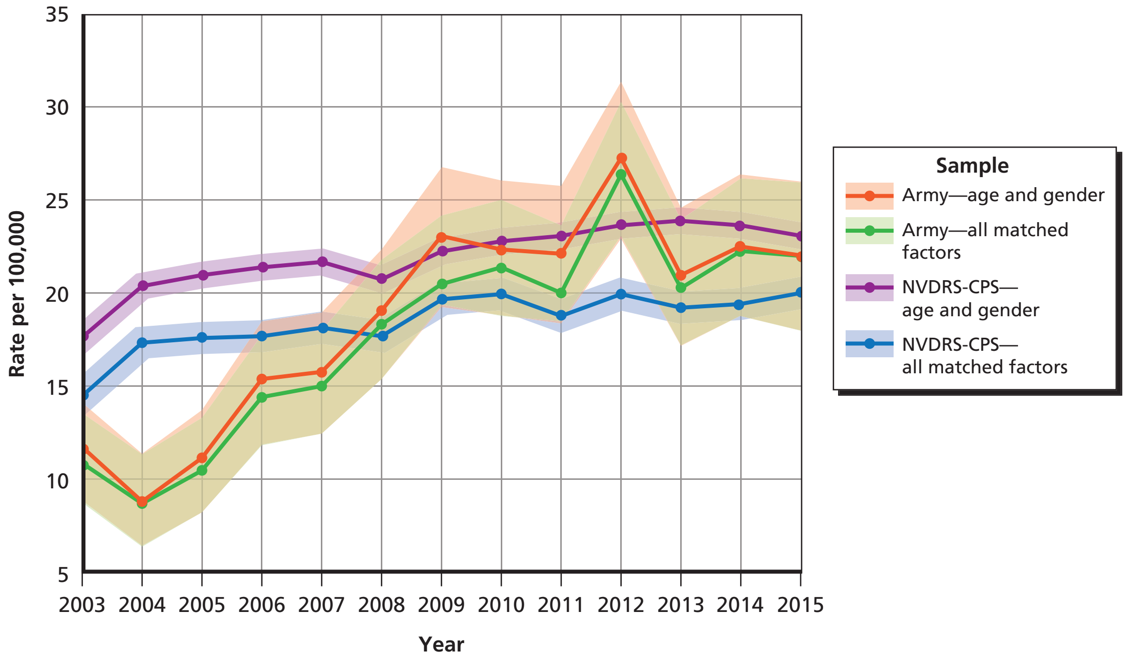
Figure F.3 shows suicide rates over time for the PS weighted Army and NVDRS-CPS samples using the PS weights that make both similar to the Army from 2015 when matching on only gender and age (red and purple lines) and when matching on all matchable factors (age, gender, race/ethnicity, education, and marital status; green and blue lines). As shown, we again see different messages when comparing suicide rates between the two populations depending on whether we match the samples on all matchable factors or just age and gender. Prior to 2008, both sets of weights show that the Army suicide rate was below that of similarly matched cases from the NVDRS-CPS sample, though the difference in suicide rates is much

Table F.1
Detailed Balance Information for Matching 2009 Army and NVDRS-CPS Samples to the 2015 Army

		2009 NVDRS-CPS Sample			2009 Army Sample		
	2015 Army	Unweighted	Weighted—All Matchable Factors	Weighted—Age + Gender	Unweighted	Weighted—All Matchable Factors	Weighted—Age + Gender
Gender							
Female	14.1%	51.7%*	14.1%	14.1%	13.2%	14.1%	14.1%
Race							
White	58.7%	71.5%*	58.7%	65.5%*	51.5%*	58.6%	51.2%*
Black	21.4%	14.1%*	21.4%	14.8%*	19.5%	21.4%	20.1%
Hispanic	12.9%	8.4%*	13.0%	12.6%	10.6%	13.0%	10.6%
Asian/Pacific Islander	5.8%	3.8%	5.8%	4.4%	15.7%*	5.8%	15.3%*
Other/unknown	1.2%	2.2%	1.2%	2.8%*	2.7%*	1.2%	2.8%*
Education level							
<BA	77.1%	62.3%*	77.0%	68.6%*	82.8%*	77.1%	81.3%*
BA	15.3%	27.1%*	15.3%	24.5%*	11.7%	15.3%	12.3%
>BA	7.7%	10.6%*	7.7%	7.0%	5.5%	7.7%	6.4%
Marital status							
Married	60.1%	55.7%*	60.1%	38.1%*	59.2%	60.1%	60.7%
Never married	34.2%	25.8%*	34.2%	54.7%*	35.0%	34.2%	33.0%
Formerly married	5.7%	18.5%*	5.7%	7.2%	5.9%	5.7%	6.3%
Age group							
Age 18–21	12.7%	5.4%*	12.7%	12.7%	14.1%	12.7%	12.7%
Age 21–24	24.3%	7.3%*	24.3%	24.3%	26.2%	24.3%	24.3%
Age 25–29	22.5%	9.5%*	22.5%	22.5%	23.7%	22.5%	22.5%
Age 30–38	25.9%	15.5%*	25.9%	25.9%	23.9%	26.0%	25.9%
Age 39–44	9.7%	11.2%*	9.7%	9.7%	8.7%	9.7%	9.7%
Age 45–54	4.6%	19.7%*	4.6%	4.6%	3.2%	4.6%	4.6%
Age 55–64	0.3%	15.3%*	0.3%	0.3%	0.2%	0.3%	0.3%
Age 65–75	0.0%	9.6%*	0.0%	0.0%	0.0%	0.0%	0.0%

* denotes when differences between the 2009 NVDRS-CPS or Army samples have an absolute effect size difference > 0.10.

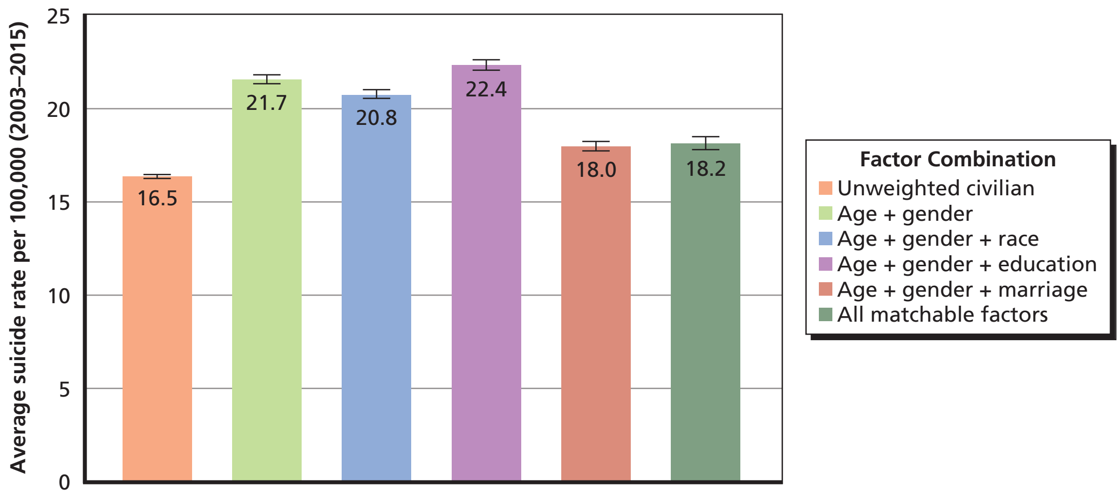
Figure F.3
Suicide Rates for Army and NVDRS-CPS Samples When Matched to 2015 Army on Only Age and Gender (red and purple lines, respectively) and All Matched Factors (green and blue lines, respectively)



smaller when matching on all matchable factors than when matching on just age and gender. In contrast, after 2008, we see when matching on just age and gender, the suicide rates in the Army and NVDRS-CPS samples look similar. In contrast, when matching on all matchable factors, the Army suicide rate more clearly surpasses the NVDRS-CPS rate after 2008. However, it still remains similar to the NVDRS-CPS rate given the overlapping confidence interval in all years, except 2012.

To gain a better understanding of which factor or factors (race/ethnicity, education, and marital status) appears to be driving the downward shift in NVDRS-CPS suicide rates after matching for more factors, we estimated three additional propensity score models for the Army 2015 analyses (just as in our primary analysis): (1) which ones matched the NVDRS-CPS and Army populations to the 2015 Army on age, gender, and race/ethnicity; (2) which matched the NVDRS-CPS and Army populations to the 2015 Army on age, gender, and education; and (3) which matched the NVDRS-CPS and Army populations to the 2015 Army on age, gender, and marital status. These findings are shown in Figure F.4. Here, we see a very strong influence of marital status on the decline in NVDRS-CPS matched suicide rates when compared with the analysis that solely matches on age and gender. On average, inclusion of marriage decreases the suicide rate by 3.7 suicides each year compared with the match which include only age and gender.

Figure F.4
Average Suicide Rate per 100,000 in the 2015 Army Matched General Population as a Function of Different Matching Factors



References

Abel, E. L., and M. L. Kruger, "Educational Attainment and Suicide Rates in the United States," *Psychological Reports*, Vol. 97, No. 1, August 2005, pp. 25–28. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/16279300

Agerbo, E., J. A. Sterne, and D. J. Gunnell, "Combining Individual and Ecological Data to Determine Compositional and Contextual Socio-Economic Risk Factors for Suicide," *Social Science & Medicine*, Vol. 64, No. 2, January 2007, pp. 451–461. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/17050054

Ahmedani, B. K., G. E. Simon, C. Stewart, A. Beck, B. E. Waitzfelder, R. Rossom, F. Lynch, A. Owen-Smith, E. M. Hunkeler, U. Whiteside, B. H. Operskalski, M. J. Coffey, and L. I. Solberg, "Health Care Contacts in the Year Before Suicide Death," *Journal of General Internal Medicine*, Vol. 29, No. 6, June 2014, pp. 870–877. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/24567199

Amini, P., H. Ahmadinia, J. Poorolajal, and M. Moqaddasi Amiri, "Evaluating the High Risk Groups for Suicide: A Comparison of Logistic Regression, Support Vector Machine, Decision Tree and Artificial Neural Network," *Iranian Journal of Public Health*, Vol. 45, No. 9, September 2016, pp. 1179–1187. As of July 26, 2019:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5149472/

Anderson S., Auquier A., Hauck W. W., et al., *Statistical Methods for Comparative Studies*, New York: John Wiley & Sons, Inc., 1980.

Anglemyer, A., M. L. Miller, S. Buttrey, and L. Whitaker, "Suicide Rates and Methods in Active Duty Military Personnel, 2005 to 2011: A Cohort Study," *Annals of Internal Medicine*, Vol. 165, No. 3, August 2, 2016, pp. 167–174. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27272476

Annesi, I., T. Moreau, and J. Lellouch, "Efficiency of the Logistic Regression and Cox Proportional Hazards Models in Longitudinal Studies," *Statistics in Medicine*, Vol. 8, No. 12, December 1989, pp. 1515–1521. As of July 26, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/2616941

Austin, Peter C., "A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications," *International Statistical Review*, Vol. 85, 2017, pp. 185–203.

Bachynski, K. E., M. Canham-Chervak, S. A. Black, E. O. Dada, A. M. Millikan, and B. H. Jones, "Mental Health Risk Factors for Suicides in the US Army, 2007–8," *Injury Prevention*, Vol. 18, No. 6, December 2012, pp. 405–412. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/22398362

Bentley, Kate H., Joseph C. Franklin, Jessica D. Ribeiro, Evan M. Kleiman, Kathryn R. Fox, and Matthew K. Nock, "Anxiety and Its Disorders as Risk Factors for Suicidal Thoughts and Behaviors: A Meta-Analytic Review," *Clinical Psychology Review*, Vol. 43, 2016, pp. 30–46. As of July 24, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/26688478

Bhatia, M. S., A. Rathi, and N. Kaur, "Confounders in Studies of Suicide by Occupation," *British Journal of Psychiatry*, Vol. 204, 2014, p. 402. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/24785769

Boffa, J. W., I. H. Stanley, M. A. Hom, A. M. Norr, T. E. Joiner, and N. B. Schmidt, "PTSD Symptoms and Suicidal Thoughts and Behaviors Among Firefighters," *Journal of Psychiatric Research*, Vol. 84, January 2017, pp. 277–283. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27810667

Boxer, P. A., C. Burnett, and N. Swanson, "Suicide and Occupation: A Review of the Literature," *Journal of Occupational and Environmental Medicine*, Vol. 37, No. 4, April 1995, pp. 442–452. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/7670900

Brenner, B., D. Cheng, S. Clark, and C. A. Camargo Jr., "Positive Association Between Altitude and Suicide in 2584 U.S. Counties," *High Altitude Medicine & Biology*, Vol. 12, No. 1, Spring 2011, pp. 31–35. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/21214344

Brent, D. A., J. A. Perper, G. Moritz, M. Baugher, J. Schweers, and C. Roth, "Suicide in Affectively Ill Adolescents: A Case-Control Study," *Journal of Affective Disorders*, Vol. 31, No. 3, July 1994, pp. 193–202. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/7963072

Briggs, Justin Thomas, and Alexander Tabarrok, "Firearms and Suicides in US States," *International Review of Law and Economics*, Vol. 37, March 2014, pp. 180–188. As of July 23, 2019:
https://www.sciencedirect.com/science/article/abs/pii/S014481881300077X

Brown, S., L. Casanave, Y. Yang, D. Haddad, and A. Stromberg, *Deaths from Violence: A Look at 18 States*, Washington, DC: National Violent Death Reporting System, 2013.

Burgette, L., D. F. McCaffrey, and B. A. Griffin, "Propensity Score Estimation with Boosted Regression," in W. Pan and H. Bai, eds., *Propensity Score Analysis: Fundamentals and Developments*, New York: Guilford Publications, Inc., 2015, pp. 49–73.

Burnham, Kenneth P., and David R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach*, 2nd ed., New York: Springer, 2002.

———, "Multimodel Inference: Understanding AIC and BIC in Model Selection," *Sociological Methods & Research*, Vol. 33, No. 2, 2004, pp. 261–304. As of July 23, 2019:
https://doi.org/10.1177/0049124104268644

Callas, Peter W., Harris Pastides, and David W. Hosmer, "Empirical Comparisons of Proportional Hazards, Poisson, and Logistic Regression Modeling of Occupational Cohort Data," *American Journal of Industrial Medicine*, Vol. 33, No. 1, 1998, pp. 33–47.

Cantor, Christopher H., and Penelope J. Slater, "Marital Breakdown, Parenthood, and Suicide," *Journal of Family Studies*, Vol. 1, No. 2, 1995, pp. 91–102.

Case, Anne, and Angus Deaton, "Rising Morbidity and Mortality in Midlife Among White Non-Hispanic Americans in the 21st Century," *Proceedings of the National Academy of Sciences*, Vol. 112, No. 49, 2015, pp. 15078–15083. As of July 23, 2019:
http://www.pnas.org/content/pnas/112/49/15078.full.pdf

CDC—*See* Centers for Disease Control and Prevention.

Centers for Disease, Control, and Prevention, "About Underlying Cause of Death, 1999–2017," CDC WONDER, Atlanta, Ga., undated-a. As of July 26, 2019:
https://wonder.cdc.gov/ucd-icd10.html

———, CDC WONDER Online Databases, Atlanta, Ga., undated-b. As of July 26, 2019:
https://wonder.cdc.gov

———, "Industry and Occupation Coding: About the NIOSH Industry & Occupation Computerized Coding System (NIOCCS)," webpage, Washington, D.C.: U.S. Department of Health and Human Services, undated-c. As of July 26, 2019:
https://www.cdc.gov/niosh/topics/coding/overview.html

———, "Regional Variations in Suicide Rates—United States, 1990–1994," *Morbidity and Mortality Weekly Report*, Vol. 46, No. 34, August 29, 1997, pp. 789–793. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/9285398

———, "Homicide and Suicide Rates—National Violent Death Reporting System, Six States, 2003," *Morbidity and Mortality Weekly Report*, Vol. 54, No. 15, April 21, 2005, pp. 377–380. As of July 23, 2019: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5415a1.htm

———, "Suicide: Facts at a Glance, 2015," Atlanta, Ga.: National Center for Injury Prevention and Control, 2015. As of July 23, 2019:
https://www.cdc.gov/violenceprevention/pdf/suicide-datasheet-a.pdf

———, "Welcome to WISQARS™ (Web-Based Injury Statistics Query and Reporting System)," database, May 3, 2018a. As of July 23, 2019:
https://www.cdc.gov/injury/wisqars/index.html

———, "Fatal Injury Data," Data and Statistics (WISQARS) database, July 12, 2018b. As of July 23, 2019: https://www.cdc.gov/injury/wisqars/fatal.html

Cheng, D., "Higher Suicide Death Rate in Rocky Mountain States and a Correlation to Altitude," *Wilderness & Environmental Medicine*, Vol. 21, No. 2, June 2010, pp. 177–178. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/20591388

Cheng, Yiling J., Edward W. Gregg, Deborah B. Rolka, and Theodore J. Thompson, "Using Multi-Year National Survey Cohorts for Period Estimates: An Application of Weighted Discrete Poisson Regression for Assessing Annual National Mortality in US Adults With and Without Diabetes, 2000–2006," *Population Health Metrics,* Vol. 14, No. 1, 2016, p. 48.

Chesney, E., G. M. Goodwin, and S. Fazel, "Risks of All-Cause and Suicide Mortality in Mental Disorders: A Meta-Review," *World Psychiatry*, Vol. 13, No. 2, June 2014, pp. 153–160. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/24890068

Christodoulou, C., A. Douzenis, F. C. Papadopoulos, A. Papadopoulou, G. Bouras, R. Gournellis, and L. Lykouras, "Suicide and Seasonality," *Acta Psychiatrica Scandinavica*, Vol. 125, No. 2, February 2012, pp. 127–146. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/21838741

Clark, T. S., and D. A. Linzer, "Should I Use Fixed or Random Effects?" *Political Science Research and Methods,* Vol. 3, No. 2, May 2015, pp. 399–408. As of July 23, 2019:
https://doi.org/10.1017/psrm.2014.32

Cohen, Jacob, "A Power Primer," *Psychological Bulletin*, Vol. 112, No. 1, July 1992, pp. 155–159. As of July 25, 2019:
https://psycnet.apa.org/buy/1992-37683-001

Conejero, I., J. Lopez-Castroman, L. Giner, and E. Baca-Garcia, "Sociodemographic Antecedent Validators of Suicidal Behavior: A Review of Recent Literature," *Current Psychiatry Reports*, Vol. 18, No. 10, October 2016, p. 94. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27595860

Conwell, Y., P. R. Duberstein, K. Connor, S. Eberly, C. Cox, and E. D. Caine, "Access to Firearms and Risk for Suicide in Middle-Aged and Older Adults," *American Journal of Geriatric Psychiatry*, Vol. 10, No. 4, July–August 2002, pp. 407–416. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/12095900

Cummings, P., T. D. Koepsell, D. C. Grossman, J. Savarino, and R. S. Thompson, "The Association Between the Purchase of a Handgun and Homicide or Suicide," *American Journal of Public Health*, Vol. 87, No. 6, June 1997, pp. 974–978. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/9224179

Curtin, S. C., M. Warner, and H. Hedegaard, "Increase in Suicide in the United States, 1999–2014," *NCHS Data Brief*, No. 241, April 2016, pp. 1–8. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27111185

Cylus, J., M. M. Glymour, and M. Avendano, "Do Generous Unemployment Benefit Programs Reduce Suicide Rates? A State Fixed-Effect Analysis Covering 1968–2008," *American Journal of Epidemiology*, Vol. 180, No. 1, July 1, 2014, pp. 45–52. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/24939978

D'Agostino, Ralph B., Mei-Ling Lee, Albert J. Belanger, L. Adrienne Cupples, Keaven Anderson, and William B. Kannel, "Relation of Pooled Logistic Regression to Time Dependent Cox Regression Analysis: The Framingham Heart Study," *Statistics in Medicine,* Vol. 9, No. 12, 1990, pp. 1501–1515.

Das, S., *High-Level Data Fusion*, Norwood, Mass.: Artech House, 2008.

Defense Manpower Data Center, DMDC Profile 2006, Arlington, Va., 2006.

Dehejia, Rajeev H., and Sadek Wahba, "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, Vol. 94, No. 448, December 1999, pp. 1053–1062. As of July 25, 2019:
http://www.uh.edu/~adkugler/Dehejia&Wahba_JASA.pdf

DMDC—*See* Defense Manpower Data Center.

DoD—*See* U.S. Department of Defense.

Efron, Bradley, "Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve," *Journal of the American Statistical Association,* Vol. 83, No. 402, 1988, pp. 414–425.

Fontanella, C. A., D. L. Hiance-Steelesmith, G. S. Phillips, J. A. Bridge, N. Lester, H. A. Sweeney, and J. V. Campo, "Widening Rural-Urban Disparities in Youth Suicides, United States, 1996–2010," *JAMA Pediatrics*, Vol. 169, No. 5, May 2015, pp. 466–473. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/25751611

Forster, Martin, and Andrew M. Jones. "The Role of Tobacco Taxes in Starting and Quitting Smoking: Duration Analysis of British Data," *Journal of the Royal Statistical Society: Series A (Statistics in Society),* Vol. 164, No. 3, 2001, pp. 517–547.

Franklin, J. C., J. D. Ribeiro, K. R. Fox, K. H. Bentley, E. M. Kleiman, X. Huang, K. M. Musacchio, A. C. Jaroszewski, B. P. Chang, and M. K. Nock, "Risk Factors for Suicidal Thoughts and Behaviors: A Meta-Analysis of 50 Years of Research," *Psychological Bulletin*, Vol. 143, No. 2, February 2017, pp. 187–232. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27841450

Gibbons, Robert D., Naihua Duan, David Meltzer, Andrew Pope, Edward D. Penhoet, Nancy N. Dubler, Charles Francis, et al., "Waiting for Organ Transplantation: Results of an Analysis by an Institute of Medicine Committee," *Biostatistics,* Vol. 4, No. 2, 2003, pp. 207–222.

Gilman, S. E., E. J. Bromet, K. L. Cox, L. J. Colpe, C. S. Fullerton, M. J. Gruber, S. G. Heeringa, L. Lewandowski-Romps, A. M. Millikan-Bell, J. A. Naifeh, M. K. Nock, M. V. Petukhova, N. A. Sampson, M. Schoenbaum, M. B. Stein, R. J. Ursano, S. Wessely, A. M. Zaslavsky, R. C. Kessler, and Starrs Collaborators Army, "Sociodemographic and Career History Predictors of Suicide Mortality in the United States Army 2004–2009," *Psychological Medicine*, Vol. 44, No. 12, September 2014, pp. 2579–2592. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/25055175

Goldsmith, S. K., T. C. Pellmar, A. M. Kleinman, W. E. Bunney, ed., *Reducing Suicide: A National Imperative*, Washington, D.C.: National Academies Press, 2002. As of July 23, 2019:
https://www.nap.edu/read/10398/chapter/1#viii

Griffin, B. A., D. McCaffrey, D. Almirall, C. Setodji, and L. Burgette, "Chasing Balance and Other Recommendations for Improving Nonparametric Propensity Score Models," *Journal of Causal Inference*, Vol. 5, No. 2, 2017. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/29503788

Hainmueller, J., "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, Vol. 20, 2012, pp. 243–263.

Haws, C. A., D. D. Gray, D. A. Yurgelun-Todd, M. Moskos, L. J. Meyer, and P. F. Renshaw, "The Possible Effect of Altitude on Regional Variation in Suicide Rates," *Medical Hypotheses*, Vol. 73, No. 4, October 2009, pp. 587–590. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/19581053

Helmkamp, J. C., "Suicides in the Military: 1980–1992," *Military Medicine*, Vol. 160, No. 2, February 1995, pp. 45–50. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/7783916

Henderson, S. N., V. B. Van Hasselt, T. J. LeDuc, and J. Couwels, "Firefighter Suicide: Understanding Cultural Challenges for Mental Health Professionals," *Professional Psychology: Research and Practice,* Vol. 47, No. 3, 2016, p. 224.

Hepner, K. A., C. P. Roth, E. M. Sloss, S. M. Paddock, P. O. Iyiewuare, M. J. Timmer, and H. A. Pincus, "Quality of Care for PTSD and Depression in the Military Health System: Final Report," *RAND Health Quarterly*, Vol. 7, No. 3, April 2018, p. 4. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/29607248

Holford, Theodore R. "The Analysis of Rates and of Survivorship Using Log-Linear Models," *Biometrics,* 1980, pp. 299–305.

Hourani, L. L., L. Davidson, M. Clinton-Sherrod, N. Patel, M. Marshall, and A. E. Crosby, "Suicide Prevention and Community-Level Indictors," *Evaluation and Program Planning*, Vol. 29, No. 4, November 2006, pp. 377–385. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/17950865

Howden, Lindsay M., and Julie A. Meyer, *Age and Sex Composition: 2010*, Washington, D.C.: 2010 Census Briefs, C2010BR-03, May 2011. As of July 23, 2019:
https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf

Hoyer, G., and E. Lund, "Suicide Among Women Related to Number of Children in Marriage," *Archives of General Psychiatry*, Vol. 50, No. 2, February 1993, pp. 134–137. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/8427553

Hsu, Chiu-Hsieh, Sylvan B. Green, and Yulei He, "A Weighted Logistic Regression Model for Estimation of Recurrence of Adenomas," *Statistics in Medicine*, Vol. 26, No. 7, March 2007, pp. 1417–1652. As of July 26, 2019:
https://doi.org/10.1002/sim.2648

Imai, K., and M. Ratkovic, "Balancing Propensity Score," *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, Vol. 76, 2014, pp. 243–263.

Ingram, D. D., and J. C. Kleinman, "Empirical Comparisons of Proportional Hazards and Logistic Regression Models," *Statistics in Medicine*, Vol. 8, No. 5, May 1989, pp. 525–538. As of July 26, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/2727473

Ingram, D. D., J. C. Kleinman, and P. Mock. "Empirical Comparisons of Proportional Hazards and Logistic Regression Models," *Statistics in Medicine,* Vol. 9, No. 4, 1990, pp. 463–464.

Ivey-Stephenson, A. Z., A. E. Crosby, S. P. D. Jack, T. Haileyesus, and M. J. Kresnow-Sedacca, "Suicide Trends Among and Within Urbanization Levels by Sex, Race/Ethnicity, Age Group, and Mechanism of Death—United States, 2001–2015," *MMWR Surveillance Summaries*, Vol. 66, No. 18, October 6, 2017, pp. 1–16. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/28981481

Johnson, F. W., P. J. Gruenewald, and L. G. Remer, "Suicide and Alcohol: Do Outlets Play a Role?" *Alcoholism: Clinical and Experimental Research*, Vol. 33, No. 12, December 2009, pp. 2124–2133. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/19764933

Judd, F., A. M. Cooper, C. Fraser, and J. Davis, "Rural Suicide—People or Place Effects?" *Australian and New Zealand Journal of Psychiatry*, Vol. 40, No. 3, March 2006, pp. 208–216. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/16476147

Kelly, T. M., and J. J. Mann, "Validity of DSM-III-R Diagnosis by Psychological Autopsy: A Comparison with Clinician Ante-Mortem Diagnosis," *Acta Psychiatrica Scandinavica,* Vol. 94, No. 5, 1996, pp. 337–343. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/9124080

Kessler, R. C., I. Hwang, C. A. Hoffmire, J. F. McCarthy, M. V. Petukhova, A. J. Rosellini, N. A. Sampson, A. L. Schneider, P. A. Bradley, I. R. Katz, C. Thompson, and R. M. Bossarte, "Developing a Practical Suicide Risk Prediction Model for Targeting High-Risk Patients in the Veterans Health Administration," *International Journal of Methods in Psychiatric Research*, Vol. 26, No. 3, September 2017. As of July 26, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/28675617

Kessler, R. C., M. B. Stein, P. D. Bliese, E. J. Bromet, W. T. Chiu, K. L. Cox, L. J. Colpe, C. S. Fullerton, S. E. Gilman, M. J. Gruber, S. G. Heeringa, L. Lewandowski-Romps, A. Millikan-Bell, J. A. Naifeh, M. K. Nock, M. V. Petukhova, A. J. Rosellini, N. A. Sampson, M. Schoenbaum, A. M. Zaslavsky, and R. J. Ursano, "Occupational Differences in US Army Suicide Rates," *Psychological Medicine*, Vol. 45, No. 15, November 2015a, pp. 3293–3304. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/26190760

Kessler, R. C., C. H. Warner, C. Ivany, M. V. Petukhova, S. Rose, E. J. Bromet, M. Brown, 3rd, T. Cai, L. J. Colpe, K. L. Cox, C. S. Fullerton, S. E. Gilman, M. J. Gruber, S. G. Heeringa, L. Lewandowski-Romps, J. Li, A. M. Millikan-Bell, J. A. Naifeh, M. K. Nock, A. J. Rosellini, N. A. Sampson, M. Schoenbaum, M. B. Stein, S. Wessely, A. M. Zaslavsky, R. J. Ursano, and Starrs Collaborators Army, "Predicting Suicides After Psychiatric Hospitalization in US Army Soldiers: The Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS)," *JAMA Psychiatry*, Vol. 72, No. 1, January 2015b, pp. 49–57. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/25390793

Klingelschmidt, J., A. Milner, I. Khireddine-Medouni, K. Witt, E. C. Alexopoulos, S. Toivanen, A. D. LaMontagne, J. F. Chastang, and I. Niedhammer, "Suicide Among Agricultural, Forestry, and Fishery Workers: A Systematic Literature Review and Meta-Analysis," *Scandinavian Journal of Work, Environment & Health*, Vol. 44, No. 1, January 1, 2018, pp. 3–15. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/29085951

Kochanek, K. D., S. L. Murphy, J. Xu, and B. Tejada-Vera, "Deaths: Final Data for 2014," *National Vital Statistics Reports*, Vol. 65, No. 4, June 2016, pp. 1–122. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27378572

Kostova, Deliana, Muhammad J. Husain, and Frank J. Chaloupka, "Effect of Cigarette Prices on Smoking Initiation and Cessation in China: a Duration Analysis," *Tobacco Control,* Vol. 26, No. 5, 2017, pp. 569–574.

Kposowa, A. J., "Association of Suicide Rates, Gun Ownership, Conservatism and Individual Suicide Risk," *Social Psychiatry and Psychiatric Epidemiology*, Vol. 48, No. 9, September 2013, pp. 1467–1479. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/23456258

Kung, H. C., J. L. Pearson, and R. Wei, "Substance Use, Firearm Availability, Depressive Symptoms, and Mental Health Service Utilization Among White and African American Suicide Decedents Aged 15 to 64 Years," *Annals of Epidemiology*, Vol. 15, No. 8, September 2005, pp. 614–621. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/16118006

Kyung-Sook, W., S. SangSoo, S. Sangjin, and S. Young-Jeon, "Marital Status Integration and Suicide: A Meta-Analysis and Meta-Regression," *Social Science & Medicine*, Vol. 197, January 2017, pp. 116–126. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/29227910

Lai, Xin, and Kelvin KW Yau. "Multilevel Mixture Cure Models with Random Effects," *Biometrical Journal,* Vol. 51, No. 3, 2009, pp. 456–466.

Laird, Nan, and Donald Olivier, "Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques," *Journal of the American Statistical Association,* Vol. 76, No. 374, 1981, pp. 231–240.

Lambert, Paul C., John R. Thompson, Claire L. Weston, and Paul W. Dickman, "Estimating and Modeling the Cure Fraction in Population-Based Cancer Survival Analysis," *Biostatistics,* Vol. 8, No. 3, 2006, pp. 576–594.

Lang, M., "The Impact of Mental Health Insurance Laws on State Suicide Rates," *Health Economics*, Vol. 22, No. 1, January 2013, pp. 73–88. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/22184054

Lavender, A., V. Ramirez-Irizarry, A. R. Bayakly, C. Koplan, and J. M. Bryan, "Violent Deaths Among Georgia Workers: An Examination of Suicides and Homicides by Occupation, 2006–2009," *American Journal of Preventive Medicine*, Vol. 51, No. 5 (Suppl. 3), November 2016, pp. S241–S250. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27745613

Lavielle, M., "Preface," in *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools*, Boca Raton, Fla.: CRC Press, 2015, pp. xv–xviii.

LeardMann, C. A., T. M. Powell, T. C. Smith, M. R. Bell, B. Smith, E. J. Boyko, T. I. Hooper, G. D. Gackstetter, M. Ghamsary, and C. W. Hoge, "Risk Factors Associated with Suicide in Current and Former US Military Personnel," *JAMA*, Vol. 310, No. 5, August 7, 2013, pp. 496–506. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/23925620

Lee, B. K., J. Lessler, and E. A. Stuart, "Improving Propensity Score Weighting Using Machine Learning," *Statistics in Medicine*, Vol. 29, No. 3, February 10, 2010, pp. 337–346. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/19960510

Lester, D., "Explaining Regional Differences in Suicide Rates," *Social Science & Medicine*, Vol. 40, No. 5, March 1995, pp. 719–721. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/7747206

Loomis, D., D. B. Richardson, and L. Elliott, "Poisson Regression Analysis of Ungrouped Data," *Occupational and Environmental Medicine,* Vol. 62, No. 5, 2005, pp. 325–329.

Lubin, G., N. Werbeloff, D. Halperin, M. Shmushkevitch, M. Weiser, and H. Y. Knobler, "Decrease in Suicide Rates After a Change of Policy Reducing Access to Firearms in Adolescents: A Naturalistic Epidemiological Study," *Suicide and Life-Threatening Behavior*, Vol. 40, No. 5, October 2010, pp. 421–424. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/21034205

Luo, Feijun, Curtis S. Florence, Myriam Quispe-Agnoli, Lijing Ouyang, and Alexander E. Crosby, "Impact of Business Cycles on US Suicide Rates, 1928–2007," *American Journal of Public Health*, Vol. 101, No. 6, June 2011, pp. 1139–1146. As of July 24, 2019: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3093269/

Luoma, J. B., C. E. Martin, and J. L. Pearson, "Contact with Mental Health and Primary Care Providers Before Suicide: A Review of the Evidence," *American Journal of Psychiatry*, Vol. 159, No. 6, June 2002, pp. 909–916. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/12042175

Mancha, Brent E., Eren Youmans Watkins, Jerrica N. Nichols, Peter G. Seguin, and Amy Millikan Bell, "Mortality Surveillance in the U.S. Army, 2005–2011," *Military Medicine,* Vol. 179, No. 12, December 2014, pp. 1478–1486. As of July 23, 2019: https://doi.org/10.7205/MILMED-D-13-00539

Martin, W. T., "Religiosity and United States Suicide Rates, 1972–1978," *Journal of Clinical Psychology*, Vol. 40, No. 5, September 1984, pp. 1166–1169. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/6333438

McCaffrey, D. F., G. Ridgeway, and A. R. Morral, "Propensity Score Estimation with Boosted Regression for Evaluating Causal Effects in Observational Studies," *Psychological Methods*, Vol. 9, No. 4, December 2004, pp. 403–425. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/15598095

Miller, M., "The Geography of Suicide," *Psychological Reports*, Vol. 47, No. 3, Part 1, December 1980, pp. 699–702. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/7220714

Miller, M., D. Azrael, and C. Barber, "Suicide Mortality in the United States: The Importance of Attending to Method in Understanding Population-Level Disparities in the Burden of Suicide," *Annual Review of Public Health*, Vol. 33, April 2012, pp. 393–408. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/22224886

Miller, M., D. Azrael, L. Hepburn, D. Hemenway, and S. J. Lippmann, "The Association Between Changes in Household Firearm Ownership and Rates of Suicide in the United States, 1981–2002," *Injury Prevention*, Vol. 12, No. 3, June 2006, pp. 178–182. As of July 23, 2019: https://www.ncbi.nlm.nih.gov/pubmed/16751449

Milner, Allison, Heidi Hjelmeland, Ella Arensman, and Diego De Leo, "Social-Environmental Factors and Suicide Mortality: A Narrative Review of over 200 Articles," *Sociology Mind*, Vol. 3, No. 2, 2013, pp. 137–148.

Morral, Andrew, "Improving Gun Policy Science," RAND.org: Gun Policy in America—Key Findings, March 2, 2018a. As of July 23, 2019:
https://www.rand.org/research/gun-policy/essays/improving-gun-policy-science.html

———, "The Science of Gun Policy: A Critical Synthesis of Research Evidence on the Effects of Gun Policies in the United States," *RAND Health Quarterly*, Vol. 8, No. 1, August 2018b, p. 5. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/30083426

National Institutes of Health, "Welcome to SOCcer," website, Washington, D.C.: U.S. Department of Health and Human Services, undated. As of July 26, 2019:
https://soccer.nci.nih.gov/

National Institute of Mental Health, "Suicide," webpage, 2016. As of July 23, 2019:
https://www.nimh.nih.gov/health/statistics/suicide.shtml

Nguyen, Thuong, Truyen Tran, Shivapratap Gopakumar, Dinh Phung, and Svetha Venkatesh, *An Evaluation of Randomized Machine Learning Methods for Redundant Data: Predicting Short and Medium-Term Suicide Risk from Administrative Records and Risk Assessments*, arXiv.org, Ithaca, N.Y.: Cornell University, May 3, 2016. As of July 26, 2019:
https://arxiv.org/abs/1605.01116

NIMH—*See* National Institute of Mental Health.

Nock, M. K., G. Borges, E. J. Bromet, C. B. Cha, R. C. Kessler, and S. Lee, "Suicide and Suicidal Behavior," *Epidemiologic Reviews*, Vol. 30, 2008, pp. 133–154. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/18653727

Nock, M. K., C. L. Dempsey, P. A. Aliaga, D. A. Brent, S. G. Heeringa, R. C. Kessler, M. B. Stein, R. J. Ursano, and D. Benedek, "Psychological Autopsy Study Comparing Suicide Decedents, Suicide Ideators, and Propensity Score Matched Controls: Results from the Study to Assess Risk and Resilience in Service Members (Army STARRS)," *Psychological Medicine*, Vol. 47, No. 15, November 2017, pp. 2663–2674. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/28502265

Nweke, N., A. Spiess, E. Corrigan, K. Kateley, T. Mitchell, K. Cevis, J. Nichols, B. E. Mancha, E. Y. Watkins, and C. Lagana-Riordan, *Surveillance of Suicidal Behavior: January through December 2014*, Gunpower, Md.: Army Public Health Center, PHR No. S.0008057-14, November 2015. As of July 24, 2019:
https://pdfs.semanticscholar.org/05fb/b6b17df1e166439a3c424fe9abd4d5fc307b.pdf

Oquendo, M. A., S. P. Ellis, S. Greenwald, K. M. Malone, M. M. Weissman, and J. J. Mann, "Ethnic and Sex Differences in Suicide Rates Relative to Major Depression in the United States," *American Journal of Psychiatry*, Vol. 158, No. 10, October 2001, pp. 1652–1658. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/11578998

Park, N., and C. Peterson, "Suicide in Happy Places Revisited: The Geographical Unit of Analysis Matters," *Applied Psychology: Health and Well-Being*, Vol. 6, No. 3, November 2014, pp. 318–323. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/24978186

Phillips, J. A., "Factors Associated with Temporal and Spatial Patterns in Suicide Rates Across U.S. States, 1976–2000," *Demography*, Vol. 50, No. 2, April 2013, pp. 591–614. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/23196429

Phillips, J. A., and K. Hempstead, "Differences in U.S. Suicide Rates by Educational Attainment, 2000–2014," *American Journal of Preventive Medicine*, Vol. 53, No. 4, October 2017, pp. e123–e130. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/28756896

Pirracchio, R., M. L. Petersen, and M. van der Laan, "Improving Propensity Score Estimators' Robustness to Model Misspecification Using Super Learner," *American Journal of Epidemiology*, Vol. 181, No. 2, January 15, 2015, pp. 108–119. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/25515168

Price, J. H., A. J. Mrdjenovich, and J. A. Dake, "Prevalence of State Firearm Mortality and Mental Health Care Resources," *Journal of Community Health*, Vol. 34, No. 5, October 2009, pp. 383–391. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/19536645

Pruitt, Larry D., Derek J. Smolenski, Nigel E. Bush, Nancy A. Skopp, Amanda Edwards-Stewart, and Timothy V. Hoyt, *Department of Defense Suicide Event Report Calendar Year 2016 Annual Report*, Washington, D.C.: U.S. Department of Defense, 2018. As of July 23, 2019:
http://www.dspo.mil/Portals/113/Documents/DoDSER%20CY%202016%20Annual%20Report_For%20Public%20Release.pdf?ver=2018-07-02-104254-717

Qin, P., and P. B. Mortensen, "The Impact of Parental Status on the Risk of Completed Suicide," *Archives of General Psychiatry*, Vol. 60, No. 8, August 2003, pp. 797–802. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/12912763

Ramchand, Rajeev, Joie D. Acosta, Rachel M. Burns, Lisa H. Jaycox, and Christopher G. Pernin, *The War Within: Preventing Suicide in the U.S. Military*, Santa Monica, Calif.: RAND Corporation, MG-953-OSD, 2011. As of July 24, 2019:
https://www.rand.org/pubs/monographs/MG953.html

Reger, M. A., D. J. Smolenski, N. A. Skopp, M. J. Metzger-Abamukang, H. K. Kang, T. A. Bullman, S. Perdue, and G. A. Gahm, "Risk of Suicide Among US Military Service Members Following Operation Enduring Freedom or Operation Iraqi Freedom Deployment and Separation from the US Military," *JAMA Psychiatry,* Vol. 72, No. 6, June 2015, pp. 561–569. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/25830941

Regoeczi, Wendy, and Duren Banks, *The Nation's Two Measures of Homicide*, Washington, D.C.: U.S. Department of Justice, Office of Justice Programs, Program Report NCJ 247060, July 2014. As of July 25, 2019:
https://www.bjs.gov/content/pub/pdf/ntmh.pdf

Reimann, Carolyn A., and Edward L. Mazuchowski, "Suicide Rates Among Active Duty Service Members Compared with Civilian Counterparts, 2005–2014," *Military Medicine*, Vol. 183, No. suppl_1, 2018, pp. 396–402. As of July 23, 2019:
http://dx.doi.org/10.1093/milmed/usx209

Reisch, T., T. Steffen, A. Habenstein, and W. Tschacher, "Change in Suicide Rates in Switzerland Before and After Firearm Restriction Resulting from the 2003 'Army XXI' Reform," *American Journal of Psychiatry*, Vol. 170, No. 9, September 2013, pp. 977–984. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/23897090

Ribeiro, J. D., P. M. Gutierrez, T. E. Joiner, R. C. Kessler, M. V. Petukhova, N. A. Sampson, M. B. Stein, R. J. Ursano, and M. K. Nock, "Health Care Contact and Suicide Risk Documentation Prior to Suicide Death: Results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS)," *Journal of Consulting and Clinical Psychology*, Vol. 85, No. 4, April 2017, pp. 403–408. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/28333538

Ridgeway, Greg, Dan McCaffrey, Andrew Morral, Lane Burgette, and Beth Ann Griffin, "Toolkit for Weighting and Analysis of Nonequivalent Groups: A Tutorial for the twang Package," July 1, 2017. As of July 23, 2019:
https://cran.r-project.org/web/packages/twang/vignettes/twang.pdf

Ridgeway, Greg, Dan McCaffrey, Andrew Morral, Beth Ann Griffin, and Lane Burgette, "Toolkit for Weighting and Analysis of Nonequivalent Groups: A Tutorial for the twang Package," 2013. As of July 23, 2019:
https://cran.r-project.org/web/packages/twang/index.html

Robinson, Cynthia, Pam Schumacher, Marie Haring Sweeney, and Jose Lainez, *Guidelines for Reporting Occupation and Industry on Death Certificates*, Cincinnati, Ohio: U.S. Dept. of Health and Human Services, Centers for Disease Control and Prevention, and National Institute for Occupational Safety and Health, 2012. As of July 23, 2019:
https://www.cdc.gov/niosh/docs/2012-149/pdfs/2012-149.pdf

Rockett, I. R., S. Wang, S. Stack, D. De Leo, J. L. Frost, A. M. Ducatman, R. L. Walker, and N. D. Kapusta, "Race/Ethnicity and Potential Suicide Misclassification: Window on a Minority Suicide Paradox?" *BMC Psychiatry*, Vol. 10, May 19, 2010, p. 35. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/20482844

Rodríguez, G., Lecture Notes on Generalized Linear Models, 2007. As of July 23 2019:
http://data.princeton.edu/wws509/notes/

Rodríguez, Germán, "Multilevel Generalized Linear Models," in *Handbook of Multilevel Analysis*, New York: Springer, pp. 335–376.

Rosenbaum, Paul R., and Donald B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, Vol. 70, No. 1, 1983, pp. 41–55.

Russ D., K. Y. Ho, J. S. Colt, K. R. Armenti, D. Baris, W. H. Chow, F. Davis, A. Johnson, M. P. Purdue, M. R. Karagas, K. Schwartz, M. Schwenn, D. T. Silverman, C. A. Johnson, and M. C. Friesen, "Computer-Based Coding of Free-Text Job Descriptions to Efficiently Identify Occupations in Epidemiological Studies," *Occupational and Environmental Medicine*, Vol. 73, No. 6, June 2016, pp. 417–424. As of July 25, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/27102331

Schmidt, Peter, and Ann Dryden Witte, "Predicting Criminal Recidivism Using 'Split Population' Survival Time Models," *Journal of Econometrics,* Vol. 40, No. 1, 1989, pp. 141–159.

Schoenbaum, M., R. C. Kessler, S. E. Gilman, L. J. Colpe, S. G. Heeringa, M. B. Stein, R. J. Ursano, and K. L. Cox, "Predictors of Suicide and Accident Death in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS): Results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS)," *JAMA Psychiatry*, Vol. 71, No. 5, May 2014, pp. 493–503. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/24590048

Seppä, Karri, Timo Hakulinen, Hyon-Jung Kim, and Esa Läärä, "Cure Fraction Model with Random Effects for Regional Variation in Cancer Survival," *Statistics in Medicine,* Vol. 29, No. 27, 2010, pp. 2781–2793.

Seppä, K., T. Hakulinen, and E. Läärä, "Avoidable Deaths and Random Variation in Patients' Survival," *British Journal of Cancer,* Vol. 106, No. 11, 2012, pp. 1846–1849.

Shrira, I., and N. Christenfeld, "Disentangling the Person and the Place as Explanations for Regional Differences in Suicide," *Suicide and Life-Threatening Behavior*, Vol. 40, No. 3, June 2010, pp. 287–297. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/20560750

Šimundić, Ana-Maria, "Measures of Diagnostic Accuracy: Basic Definitions," *Medical and Biological Sciences*, Vol. 22, No. 4, 2008, pp. 61–65. As of July 23, 2019:
https://www.cm.umk.pl/en/images/medical/medical_22_4.pdf

Smith, N. D., and I. Kawachi, "State-Level Social Capital and Suicide Mortality in the 50 U.S. States," *Social Science & Medicine*, Vol. 120, November 2014, pp. 269–277. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/25262314

Srinivasan, Jayakanth, Millard D. Brown, Christopher G. Ivany, and Jonathan Woodson, "How the U.S. Army Personalized Its Mental Health Care," *Harvard Business Review*, December 7, 2016. As of July 23, 2019:
https://hbr.org/2016/12/how-the-u-s-army-personalized-its-mental-health-care

Stack, S., "Occupation and Suicide," *Social Science Quarterly,* Vol. 82, No. 2, 2001, pp. 384–396.

Stallones, L., "Suicide and Potential Occupational Exposure to Pesticides, Colorado 1990–1999," *Journal of Agromedicine*, Vol. 11, No. 3–4, 2006, pp. 107–112. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/19274902

Stallones, L., T. Doenges, B. J. Dik, and M. A. Valley, "Occupation and Suicide: Colorado, 2004–2006," *American Journal of Industrial Medicine*, Vol. 56, No. 11, November 2013, pp. 1290–1295. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/23873359

Stone, D. M., T. R. Simon, K. A. Fowler, S. R. Kegler, K. Yuan, K. M. Holland, A. Z. Ivey-Stephenson, and A. E. Crosby, "Vital Signs: Trends in State Suicide Rates—United States, 1999–2016 and Circumstances Contributing to Suicide—27 States, 2015," *Morbidity and Mortality Weekly Report*, Vol. 67, No. 22, June 8, 2018, pp. 617–624. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/29879094

Swaim, Paul, and Michael Podgursky, "Female Labor Supply Following Displacement: A Split-Population Model of Labor Force Participation and Job Search," *Journal of Labor Economics,* Vol. 12, No. 4, 1994, pp. 640–656.

Tiesman, H. M., S. Konda, D. Hartley, C. Chaumont Menendez, M. Ridenour, and S. Hendricks, "Suicide in U.S. Workplaces, 2003–2010: A Comparison with Non-Workplace Suicides," *American Journal of Preventive Medicine*, Vol. 48, No. 6, June 2015, pp. 674–682. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/25794471

Tondo, L., M. J. Albert, and R. J. Baldessarini, "Suicide Rates in Relation to Health Care Access in the United States: An Ecological Study," *Journal of Clinical Psychiatry*, Vol. 67, No. 4, April 2006, pp. 517–523. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/16669716

Trofimovich, L., M. A. Reger, D. D. Luxton, and L. A. Oetjen-Gerdes, "Suicide Risk by Military Occupation in the DoD Active Component Population," *Suicide and Life-Threatening Behavior*, Vol. 43, No. 3, June 2013, pp. 274–278. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/23347281

U.S. Army, "U.S. Army Demographics: FY16 Army Profile," Washington, D.C., 2016a. As of July 23, 2019:
https://m.goarmy.com/content/dam/goarmy/downloaded_assets/pdfs/advocates-demographics.pdf

———, "U.S. Army Ready and Resilient," webpage, 2016b. As of July 23, 2019:
http://readyandresilient.army.mil/

U.S. Census Bureau, "Quick Facts," webpage, undated. As of July 23, 2019:
https://www.census.gov/quickfacts/fact/table/US/PST045217

———, "Selected Social Characteristics in the United States: 2012–2016 American Community Survey 5-Year Estimates—2016," American FactFinder, 2016. As of July 23, 2019:
https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_DP02&src=pt

———, "America's Families and Living Arrangements: 2017" (data tables), in *Current Population Survey, 2017 Annual Social and Economic Supplement*, Washington, D.C.: U.S. Census Bureau, May 8, 2017a. As of July 23, 2019:
https://www.census.gov/data/tables/2017/demo/families/cps-2017.html

———, "Educational Attainment in the United States: 2017," in *Current Population Survey, 2017 Annual Social and Economic Supplement*, Washington, D.C.: U.S. Census Bureau, December 14, 2017b. As of July 23, 2019:
https://www.census.gov/data/tables/2017/demo/education-attainment/cps-detailed-tables.html

U.S. Department of Defense, *2016 Demographics Report: Profile of the Military Community*, Washington, D.C.: U.S. Department of Defense, 2016. As of July 23, 2019:
http://download.militaryonesource.mil/12038/MOS/Reports/2016-Demographics-Report.pdf

———, "STARRS-LS: Study to Assess Risk & Resilience in Service Members—Longitudinal Study," website, 2018. As of July 23, 2019:
http://starrs-ls.org/#/page/army-starrs-studies

van der Laan, M. J., E. C. Polley, and A. E. Hubbard, "Super Learner," *Statistical Applications in Genetics and Molecular Biology*, Vol. 6, 2007, p. Article25. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/17910531

Veevers, J. E., "Parenthood and Suicide: An Examination of a Neglected Variable," *Social Science & Medicine*, Vol. 7, No. 2, February 1973, pp. 135–144. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/4695265

Verplancke, T., S. Van Looy, D. Benoit, S. Vansteelandt, P. Depuydt, F. De Turck, and J. Decruyenaere, "Support Vector Machine Versus Logistic Regression Modeling for Prediction of Hospital Mortality in Critically Ill Patients with Haematological Malignancies," *BMC Medical Informatics and Decision Making*, Vol. 8, Article 56, 2008. As of July 26, 2019:
https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-8-56

Violanti, J. M., A. Mnatsakanova, C. M. Burchfiel, T. A. Hartley, and M. E. Andrew, "Police Suicide in Small Departments: A Comparative Analysis," *International Journal of Emergency Mental Health*, Vol. 14, No. 3, 2012, pp. 157–162. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/23894796

Walsh, Colin G., Jessica D. Ribeiro, and Joseph C. Franklin, "Predicting Risk of Suicide Attempts over Time Through Machine Learning," *Clinical Psychological Science*, April 11, 2017. As of July 26, 2019:
https://doi.org/10.1177/2167702617691560

Wang, P. S., M. Lane, M. Olfson, H. A. Pincus, K. B. Wells, and R. C. Kessler, "Twelve-Month Use of Mental Health Services in the United States: Results from the National Comorbidity Survey Replication," *Archives of General Psychiatry*, Vol. 62, No. 6, June 2005, pp. 629–640. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/15939840

Watkins, E. Y., A. Spiess, I. Abdul-Rahman, C. Hill, N. Gibson, J. Nichols, V. McLeod, L. Johnson, T. Mitchell, J. A. Pecko, and K. Cox, "Adjusting Suicide Rates in a Military Population: Methods to Determine the Appropriate Standard Population," *American Journal of Public Health*, Vol. 108, No. 6, June 2018, pp. 769–776.
https://www.ncbi.nlm.nih.gov/pubmed/29672151

Wenger, Jeffrey B, Ellen M. Pint, Tepring Piquado, Michael G. Shanley, Trinidad Beleche, Melissa A. Bradley, Jonathan Welch, Laura Werber, Cate Yoon, Eric J. Duckworth, and Nicole H. Curtis, *Helping Soldiers Leverage Army Knowledge, Skills, and Abilities in Civilian Jobs*, Santa Monica, Calif.: RAND Corporation, RR-1719-A, 2017. As of July 23, 2019:
https://www.rand.org/pubs/research_reports/RR1719.html

Westreich, D., J. Lessler, and M. J. Funk, "Propensity Score Estimation: Neural Networks, Support Vector Machines, Decision Trees (CART), and Meta-Classifiers as Alternatives to Logistic Regression," *Journal of Clinical Epidemiology*, Vol. 63, No. 8, August 2010, pp. 826–833. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/20630332

Willett, John B., and Judith D. Singer, "Investigating Onset, Cessation, Relapse, and Recovery: Why You Should, and How You Can, Use Discrete-Time Survival Analysis to Examine Event Occurrence," *Journal of Consulting and Clinical Psychology,* Vol. 61, No. 6, 1993, p. 952.

Wintemute, G. J., C. A. Parham, J. J. Beaumont, M. Wright, and C. Drake, "Mortality Among Recent Purchasers of Handguns," *New England Journal of Medicine*, Vol. 341, No. 21, November 18, 1999, pp. 1583–1589. As of July 23, 2019:
https://www.ncbi.nlm.nih.gov/pubmed/10564689

Wooldridge, J., *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Mass.: MIT Press, 2002.

Xu, Ying, K. F. Lam, Feifei Zhou, Paul SF Yip, and Ray Watson, "A Study of Suicide Risk Using a Cox Cure Model via a Retrospective Sampling and Multiple Imputation," *Communications in Statistics-Theory and Methods,* Vol. 41, No. 18, 2012, pp. 3389–3402.

Zhang, Mei-Jie, Xu Zhang, and Thomas H. Scheike, "Modeling Cumulative Incidence Function for Competing Risks Data," *Expert Review of Clinical Pharmacology,* Vol. 1, No. 3, 2008, 391–400.

Zou, Guangyong, "A Modified Poisson Regression Approach to Prospective Studies with Binary Data," *American Journal of Epidemiology,* Vol. 159, No. 7, 2004, pp. 702–706.

Zou, G. Y., and Allan Donner, "Extension of the Modified Poisson Regression Model to Prospective Studies with Correlated Binary Data," *Statistical Methods in Medical Research,* Vol. 22, No. 6, 2013, pp. 661–670.

Zubizarreta, J. R., "Stable Weights That Balance Covariates for Estimation with Incomplete Outcome Data," *Journal of the American Statistical Association*, Vol. 110, No. 511, 2015, pp. 910–922. As of July 23, 2019:
https://doi.org/10.1080/01621459.2015.1023805

RAND ARROYO CENTER

Over the past 15 years, the suicide rate among members of the U.S. armed forces has doubled, with the greatest increase observed among soldiers in the Army. This increasing rate is paralleled by a smaller increase in the general U.S. population, observed across both genders, in virtually every age group and in nearly every state. An empirical question exists: What is the extent or degree to which the suicide trend in the Army is unique to that service, relative to what is observed in the general U.S. population?

The Army has typically attempted to address this question by standardizing the general population to look like the Army on demographic characteristics. However, given the rise in suicide rates over the past decade, the Army wanted to better understand whether standardization based solely on age and gender is enough. Expanding the characteristics on which the general population is standardized to match the Army could be useful to gain a better understanding of the suicide trends in the Army. However, such a change also brings with it some challenges, including the lack of readily available data in the general U.S. population. In addition, even an expanded set of characteristics still results in having a large number of unmeasured factors that cannot be included in this type of analysis.

In this report, the authors explore how accounting for age, gender, race/ethnicity, time, marital status, and educational attainment affects suicide rate differences between soldiers and a comparable subset of the general U.S. population.

$37.00

ISBN-10 1-9774-0359-X
ISBN-13 978-1-9774-0359-9
53700
9 781977 403599

www.rand.org

RR-3025-A