The Effects of Vietnam-Era Military Service on the Long-Term Health of Veterans: A Bounds Analysis *

Xintong Wang† Carlos A. Flores‡ Alfonso Flores-Lagunes§

October 26, 2021

Abstract

We analyze short- and long-term health effects of U.S. Vietnam-era military service on the population of veterans using a restricted version of the National Health Interview Survey and employing the draft lotteries as an instrumental variable. We start by assessing whether the instrument satisfies the exclusion restriction. Since we do not find evidence against its validity, we conduct inference on the health effects of military service for individuals who comply with the draft-lotteries assignment (the "compliers"). We do not find evidence of detrimental health effects on compliers, consistent with prior literature. For those who volunteer for enlistment (the "always takers"), estimated bounds based on a mild mean weak monotonicity assumption indicate statistically significant detrimental health effects that appear 20 years after the end of the conflict. As a group, veterans experience similar detrimental health effects from military service. Results have implications for compensation and health care policies of veterans.

Key words and phrases: Veteran health; Treatment effects; Bounds; Instrumental variables

JEL classification: I12, I18, C31, C36

* We thank Carlos Dobkin for his assistance in the application for the restricted-use data. We particularly thank Adrienne Jones and the National Center for Health Statistics, the U.S. Census Bureau, and the Federal Statistics Research Data Center at Cornell for all the gracious support given to the access to the restricted data used in this paper. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the Research Data Center, the National Center for Health Statistics, or the Centers for Disease Control and Prevention. Flores gratefully acknowledges support from the Orfalea College of Business at California Polytechnic State University. Flores-Lagunes gratefully acknowledges support from the Appleby-Mosher fund at the Maxwell School and the services and support of the Center for Aging and Policy Studies, funded by NIA Center Grant P30AG066583. The Online Appendix to this paper can be found at the following link.

† Department of Accounting, Economics, and Finance, Slippery Rock University of Pennsylvania, Slippery Rock, PA; Telephone: +1 724-738-2579; Email: xintong.wang@sru.edu.

‡ Orfalea College of Business, California Polytechnic State University, 1 Grand Avenue, San Luis Obispo, CA 93407; Telephone: +1 805-756-2907; Email: cflore32@calpoly.edu.

§ **Corresponding author**. Department of Economics and Center for Policy Research, Syracuse University; IZA, and GLO; 426 Eggers Hall, Syracuse NY 13244-1020; Telephone: +1 315-443-9045; Email: afloresl@maxwell.syr.edu.

1 Introduction

The U.S. veteran population is on average less healthy than the non-veteran population, for instance, in terms of the prevalence of chronic conditions (Kramarow and Pastor, 2012), psychological distress (Kramarow and Pastor, 2012), work limitations (Kramarow and Pastor, 2012, Gustman et al. 2016), and self-reported health (Teachman, 2011, Gustman et al. 2016). At the same time, the number of veterans who receive disability compensation increased by 55% (from 2.3 to 3.5 million) from 2000 to 2013, despite a decrease of 17% in the total living veteran population (Congressional Budget Office, 2014). In particular, veterans who served during the Vietnam era received the highest average annualized payments from Veterans Affairs (VA) Disability Compensation (VDC) among all veterans, including those who served in World War II and the Korea War (Congressional Budget Office, 2014). While Vietnam-era veterans are on average less healthy than non-veterans, the corresponding literature—which largely focuses on a subset of veterans—finds inconclusive causal statistical evidence on the effect of Vietnam-era military service on veterans' health.

There is a large literature that estimates the causal effect of Vietnam-era military service on the health outcomes and behaviors of veterans. To estimate this causal effect, researchers have to overcome the selection problem: people with certain pre-induction characteristics related to their health may be more inclined to join the military (Seltzer and Jablon, 1974, Angrist, 1991, Watkins and Sherk, 2008, Carter et al. 2017). A considerable number of studies attempt to overcome this selection problem by exploiting variation in the military service status induced by a randomly assigned instrumental variable (IV). A widely used IV is the U.S. Vietnam-era draft lotteries that determined induction eligibility solely based on random sequence numbers (RSNs) assigned to males' birthdays. Most—if not all—of the studies using this IV undertake inference on the causal health effects of military service on the subpopulation whose military service was induced by the draft lotteries, a subpopulation known as the "compliers" (Angrist et al., 1996).

Several papers that use the U.S. Vietnam-era draft lotteries as an IV for military service find weak or no evidence that military service affects health outcomes such as disability rates (Angrist et al., 2010; Davies et al., 2015), self-reported health, activity limitations, chronic conditions, and mental health (Dobkin and Shabani, 2009). Some exceptions pertain to smoking and its consequences. Eisenberg and Rowe (2009) find that Vietnam-era military service statistically increases smoking among veterans during 1978-1980, although the estimated effect became statistically insignificant by 1997-2005. Schmitz and Conley (2016) find that Vietnam-era military service statistically increased smoking and the risk of being diagnosed with cancer or hypertension at older ages, but this effect was only statistically significant among veterans who have a high genetic predisposition to smoking. Another exception pertains to mortality, where the early study by Hearst et al. (1986) found that Vietnam-era military service increased suicides and death from motor vehicle accidents. Angrist et al. (1996), however, did not find a statistically significant effect on suicides or overall mortality using the same approach but somewhat different sample.[1]

We analyze the effects of military service on a comprehensive list of health outcomes and behaviors using a restricted version of the National Health Interview Surveys (NHIS) 1974-2013. The NHIS is the main cross-sectional data source on health that is representative of the civilian non-institutionalized population of the United States. These data allow us to construct the individual's draft eligibility status from the Vietnam-era draft lotteries, which we use as an IV, and to examine the evolution of the health effects of military service over time, including a measure of mortality, from the end of the Vietnam War to 40 years out.[2] For this purpose, we adopt the different NHIS survey years of 1974-1981, 1982-1996, 1997-2005, and 2006-2013. We focus on health outcomes typically analyzed in the literature on veterans, such as activity and work limitations and self-reported health. We also analyze a number of activity-limiting chronic conditions, other chronic conditions, and the health behaviors of smoking and drinking.

A key distinctive feature of our analysis is that, while we use the draft lotteries as an IV, we go beyond the estimation of the health effects of military service for compliers and estimate bounds on the health effects for volunteers or "always takers" (those who are veterans regardless of draft eligibility), who comprise most of the veteran population. Together, the subpopulations of compliers and volunteers make up the group of Vietnam-era veterans (i.e., the "treated group"). Thus, our results apply to what is, arguably, the main population of interest. Additionally, the estimated bounds on the health effects of military service for volunteers are likely more informative about the corresponding effects for the current U.S. all-volunteer forces than the estimated health effects for compliers.

Our analysis starts by assessing the validity of the draft lotteries IV. Specifically, we assess whether the draft lotteries IV has a net or direct effect on the health outcomes and behaviors we analyze that is independent (or net) of military service, thereby assessing an implication of the requirement that the IV affects the outcome only through its effect on the treatment variable (Angrsit et al., 1996). Several studies (e.g., Card and Lemieux, 2001; Bailey and Chyn, 2020) have raised the possibility that, due to draft avoidance behaviors, the Vietnam-era draft lotteries IV may have a net effect on various outcomes such as college attendance, fertility, and incarcerations.[3] Therefore, we bound the average net effect of the draft eligibility

[1]Using a similar military draft lotteries event in Australia during the Vietnam War as the IV for military service, Siminski and Ville (2011) find no statistical evidence of elevated mortality between 1994 and 2007. In contrast, also for Australia, Siminski and Ville (2012) find that serving during the Vietnam War increased the self-report of disability in 2006 and the receipt of the veterans' disability compensation in 2009.

[2]The mortality outcome considered here comes from a restricted-use merge of mortality records to the NHIS data produced by the National Center on Health Statistics (NCHS), which is available only for a subset of years.

[3]Examples of draft-avoidance behaviors are college enrollment, migration out of the U.S., fatherhood, and

status (the IV) on the health outcomes and behaviors for the subpopulation of "never takers" (those who are non-veterans regardless of draft eligibility) by employing the bounds proposed by Flores and Flores-Lagunes (2013) and Chen et al. (2020); see also the recent application in Wang and Flores-Lagunes (2020).

Since we do not find evidence of a non-zero net or direct effect, we regard the draft lotteries IV as valid and move on to report point estimates of the effects of military service on health outcomes and behaviors for the subpopulation of compliers. These point estimates—which apply to less than 25% of the Vietnam-era veteran population—are obtained using IV estimation (Imbens and Angrist, 1994; Angrist et al., 1996). They update the corresponding estimates obtained in a similar way in Dobkin and Shabani (2009) and Eisenberg and Rowe (2009), by using the latest NHIS surveys. Lastly, we estimate bounds on the health effects of military service for the rest of the Vietnam-era veteran population, which consists of individuals who volunteered for enlistment, and who represent over 75% of the population of veterans in our sample. Our estimates employ the bounds proposed in Chen et al. (2018), which are derived under the standard IV assumptions. In addition, these bounds rely on a relatively mild mean weak monotonicity assumption that relates the average potential outcomes of the volunteers to those of the never takers. Specifically, we pose that, in the absence of military service for both subpopulations, the average potential health outcomes of the never takers are no better than those of the volunteers. As we discuss in detail later, the physical, medical, and mental examinations by the U.S. military to determine prospective recruits' qualifications for military service, as well as other considerations, help to justify the plausibility of this mean weak monotonicity assumption. In other words, this mean weak monotonicity assumption statistically formalizes the accepted notion of positive self-selection into the military based on health (e.g., Seltzer and Jablon, 1974; Bedard and Deschênes, 2006; Eisenberg and Rowe, 2009), which allows us to estimate informative bounds on the effects of military service on health outcomes of volunteers and all veterans.

The key findings are as follows. First, we find no statistical evidence of violations of the exclusion restriction for the subpopulation of never takers: the estimated bounds on the average net effects of the draft lotteries IV do not rule out zero for any of the health outcomes and behaviors we consider. This is consistent with and complements similar evidence in Kitagawa (2015) and Mourifié and Wan (2017), who assessed the validity of the same IV in the context of labor earnings outcomes. Second, our point estimates of the average effect of Vietnam-era military service on the health outcomes and behaviors of complier veterans do not provide consistent evidence of statistically significant effects. The majority of the estimated effects of military service are statistically insignificant. The lack of clarity in the results for compliers appears in line with prior literature (e.g., Dobkin and Shabani, 2009; Angrist et al.,

failure of the "moral standards" of the military (e.g., Suttler, 1970; Shapiro and Striker, 1970; Baskir and Strauss, 1978).

2010; Davies et al., 2015) that largely finds weak or no health effects of Vietnam-era military service.

Third, for Vietnam-era volunteers, we document that their estimated bounds show clear statistically significant detrimental health effects that seem to appear over time. In the approximately eight years after the end of the conflict (the 1974-1981 survey period), no military service effects are statistically significant (i.e., the estimated bounds and their confidence intervals do not exclude zero) for any health outcomes, with the exception of smoking. The estimated bounds, separately estimated for whites and nonwhites, are consistent with an effect on smoking of at least about 13.8 pp (44.9%) and at most about 46.5 pp (151.2%). By the second survey period—nine to twenty-three years after the end of the conflict—the health of white volunteers shows statistically significant detrimental impacts of military service on activity limitation, work limitation, self-reported fair/poor health, muscoskeletal conditions, and smoking. For nonwhite volunteers, only the detrimental military service effect on smoking is statistically significant. To illustrate the magnitude of the effects, consider the health outcome of activity limitation. The estimated lower bound for white volunteers in the second survey period indicates that military service increases activity limitation by at least 1.1 pp (9.2% relative to the nonveteran mean), while for nonwhites the effect is at least 1.8 pp, although the 95% confidence intervals (marginally) include zero for the latter group.

For the last two survey periods of 1997-2005 and 2006-2013—twenty-four to forty years after the end of the conflict—we find statistically significant detrimental military service effects for both white and nonwhite volunteers on a considerable number of health outcomes. To name a few, for both white and nonwhite volunteers, military service statistically increases the incidence of activity limitation, work limitation, and self-reported fair/poor health. Military service also statistically increases their incidence on a myriad of other chronic conditions, for example, back/neck conditions, lower back conditions, joint conditions, depression, and diabetes. As for health behaviors, smoking continues to exhibit a statistically significant effect, and drinking is statistically significant as well. Using as an illustration again the health outcome activity limitation, the estimated lower bounds for white volunteers indicate that military service increases activity limitation by at least 7.1 pp (64.2%) and 5.8 pp (50.3%) in the last two survey periods, respectively. For nonwhites, their estimated lower bounds indicate that the same effects are at least 4.3 pp (40.2%) and 11.7 pp (110.9%).[4]

Fourth, perhaps not surprisingly given the previous results for volunteers and the fact that they represent over 75% of the Vietnam-era veterans, the military service effects on the entire population of veterans follow similar trends to those of the volunteers. We also document results on the effect of Vietnam-era military service on the mortality of veterans by 2011, and discuss their implications for the previous findings about the military service impact on health

[4]Throughout our analysis, we employ a conservative multiple testing procedure to adjust the statistical inference on families of multiple health outcomes and behaviors, as discussed later.

outcomes and behaviors. Finally, an analysis of the educational attainment of compliers and volunteers prior to the Vietnam-era draft lotteries suggests a possible factor contributing to their disparate military service effects. Draft compliers were significantly more likely to have graduated from high school at baseline relative to volunteers. Assuming that veterans with pre-draft completed high school were more likely to take advantage of educational benefits through the GI Bill, it is possible that the long-run health of complier veterans may have benefited from the protective effect of education (e.g., Oreopoulos and Salvanes, 2011) relative to that of volunteers.

Our findings offer several potential policy implications. In general, our results suggest that there have been detrimental health effects of Vietnam-era military service, which show up and appear to worsen over time. They also suggest that these detrimental effects are primarily driven by males who volunteered for enlistment, and not by compliers. These results have relevant potential implications for policies regarding compensation of veterans after service. Previous studies have shown that Vietnam-era military service increases federal transferred income (e.g., Angrist et al., 2010). Some studies have attributed the recent rise in VDC compensation costs to a policy change in 2001 that allowed "boots on the ground" Vietnam veterans with diabetes mellitus to qualify for VDC benefits (Singleton, 2009, Autor et al., 2011), and the increased post-traumatic stress disorder (PTSD) diagnosis among Vietnam veterans (Autor et al., 2011). By documenting that the Vietnam-era military service statistically significantly worsened veteran's general health outcomes, depression (which is highly intercorrelated and comorbid to PTSD), and diabetes, among others, our results suggest that the increased VDC compensation could had been induced by the Vietnam-era veterans who suffer from these chronic conditions and limitations. Lastly, our results showing the short- and long-term smoking inducing effect of military service among Vietnam-era veterans (driven by volunteers), suggest that recent policies of increasing cigarette prices on military bases and an increase of minimum smoking age may help to improve veterans' healthy behaviors and reduce the healthcare cost for veterans in the long-run.

2 The U.S. Vietnam-War Draft Lotteries

On December 1, 1969, the Selective Service System conducted the first Vietnam-War lottery drawing to determine the order in which men born in the years of 1944-1950 were called to report for possible induction into the military from the beginning of 1970. In the televised draft lotteries, 366 capsules containing birth dates were placed into a glass container. The capsules were drawn by hand and assigned to a random sequence number (RSN) that ranged from 1 to 366. The local draft boards continued to draft men into military service in the order of the RSNs until manpower requirements were met. Males who received lower numbers for their birth dates had a higher probability of being called to serve. The Selective

Service System continued the draft lotteries again in 1970 and 1971, covering males born in 1951 and 1952 respectively. The last lottery numbers called became the ex-post draft eligibility cut-offs, which correspond to 195 for the 1969 draft lotteries, 125 for the 1970 draft lotteries and 95 for the 1971 draft lotteries.

Receiving a lottery number below the cutoff did not equate to subsequent induction into the military. On the one hand, males could volunteer to serve when their lottery numbers had not been called. On the other hand, draft-eligible males were subjected to physical, medical, and mental examinations to determine their qualifications for military service—only males who were classified as Class I-A (ready for military service) and I-A-O (conscientious objector available for noncombatant military service only) by the local draft board could be drafted (Shapiro and Striker, 1970). Furthermore, conscription exemptions such as paternal deferment and "conscientious objector" status were available during the years of the Vietnam-era draft lotteries (Baskir and Strauss, 1978), while college deferments exemptions were available until 1971 (Card and Lemieux, 2001). Studies have suggested that these deferments might have led to behavioral responses among males subject to the draft, such as increased male college attendance (Card and Lemieux, 2001) and draft induced marriage (Hanson, 2013). Meanwhile, other behavioral responses undertaken by males did not pertain to legal conscription exemptions (Shapiro and Striker, 1970), such as those intended to purposefully fail the moral standards of the military (e.g., committing crimes).

3 Data

The data source in this paper is a restricted version of the National Health Interview Survey (NHIS) 1974-2013. NHIS is the principal cross-sectional data source on health that is representative of the civilian non-institutionalized population of the United States. Due to changes in the survey design of NHIS, we separately analyze the survey periods of 1974-1981, 1982-1996, 1997-2005, and 2006-2013. This separate analysis allows us to capture the evolution over time of the health effects of the Vietnam-era military service over the short-run of up to 8 years after the end of the conflict (1974-1981), 9 to 23 years (1982-1996), 24 to 32 years (1997-2005), and 33 to 40 years (2006-2013). These data are similar to those used in Dobkin and Shabani (2009), who used restricted NHIS 1974-2004 data, and to Eisenberg and Rowe (2009), who used restricted NHIS 1978-1980 and 1997-2005 data.

We follow the large literature on economic analyses of the draft lotteries (e.g., Angrist et al., 2010) and focus on the 1948-1952 birth cohorts.[5] We employ the restricted-use birth

[5]A reason for the previous studies leaving out males from the 1944-1947 birth cohorts is that the effect of the draft eligibility on military service for them is small (e.g., Angrist et al., 2010). Another reason is that many of the males born in 1944-1947 had been subjected to the local drafts during the Vietnam era when they were between the ages of 18 ½ and 25, before the national lottery draft was implemented. Thus, omitting

date of each respondent in the NHIS to construct the Vietnam-era draft lotteries eligibility. The RSNs from 1 to 366 of each respondent is constructed based on the exact birth date information in the NHIS and the lottery numbers obtained from the Selective Service System website. We then define a binary draft-eligibility variable with the value of 1 (draft-eligible) if an individual received a lottery number under the draft-eligibility cutoff and 0 (draft-ineligible) if he received a lottery number above the draft-eligibility cutoff.

The definition of the military service status variable varies over time due to a change in the question that forms the basis for this definition. Our definition of this key variable follows prior studies using the NHIS data (e.g., Dobkin and Shabani, 2009; Eisenberg and Rowe, 2009). For surveys prior to 1997, the military service indicator variable is constructed based on whether an individual ever served on active duty in the Armed Forces of the United States during the Vietnam era. For surveys after 1996, the military service indicator variable is based on whether a person has ever been honorably discharged from active duty in the U.S. Army, Navy, Air Force, Marine Corps, or Coast Guard. Thus, this variable excludes individuals who served but did not receive an honorable discharge from the military. While not ideal, the "honorable discharge" variable is the only variable in the NHIS that measures military service during the survey periods after 1996, and thus has been employed by studies using those survey periods.

A key advantage of using the NHIS data is that it contains a rich list of health outcomes and behaviors. We classify these outcomes into four categories: general health outcomes, health behaviors, activity-limiting chronic conditions, and other chronic conditions. It is important to keep in mind that some of the specific health outcomes within these categories are not available in every one of the four survey periods in which we divide the NHIS 1974-2016 because of changes in the questions asked by the survey. The general health outcomes correspond to self-reported fair or poor health, activity limitation, activity unable, work limitation, and work unable.[6] These are health outcomes that have been the primary focus of earlier studies (e.g., Dobkin and Shabani, 2009), and they represent measures of overall health. Fair or poor health is coded with value 1 for those who reported "fair/poor" health on the 5-point health scale of "excellent/very good/good/fair/poor", and is coded as 0 otherwise. Studies have shown that self-reported health is a strong predictor of mortality (Pietz and Petersen, 2007; Schnittker and Bacak, 2014). Activity limitation measures whether a person is limited in any kind or amount in major activities, or is unable to perform major activities. Activity unable measures whether a person is unable to perform major activities. The work limitation variable measures whether a person is limited in the kind or amount of work because of a physical, mental or emotional problem, or a person is unable to perform any work. The work

these birth cohorts avoids potential contamination from the effects of the local drafts.

[6]The activity unable variable is only available in survey periods of 1974-1981 and 1982-1996; while the work limitation and work unable variables are not available in the survey period of 1974-1981.

unable variable measures whether a person is unable to perform any work. Work and activity limitation outcomes are relevant as they are directly related to males' labor force participation and social benefits claims, such as the VDC.

We concentrate on two health behaviors: current smoker and current drinker. While not directly health outcomes, these behaviors have been widely linked to unfavorable health outcomes (e.g., U.S. Department of Health and Human Services, 2014; Lown et al., 2007). The health outcomes under activity-limiting chronic conditions correspond to the answers to the follow-up questions in the NHIS about which chronic conditions cause the activity limitation. We consider over thirteen of this type of chronic conditions, depending on the specific survey years considered. Lastly, the health outcomes under other chronic conditions correspond to the answers to explicit survey questions about whether the individual has experienced a given condition in the past 12 months. The difference with the previous group of conditions is that these do not necessarily cause activity limitations. We consider twenty-three of these chronic conditions, which are available only in the 1997-2005 and 2006-2013 survey years.

Table 1 presents summary statistics for the general health outcomes and health behaviors by veteran status and white/nonwhite for the different years of NHIS surveys. Regarding the draft-eligibility status, as expected, there are more draft-eligible males among the veterans than among the nonveterans, and the proportions of draft-eligible males are similar over the four survey periods. In the first two survey periods (1974-1981 and 1982-1996), the average self-reported poor or fair health indicator shows no statistically significant difference between veterans and nonveterans. This later changes as veterans show a statistically significantly worse average self-reported health than nonveterans for whites in 1997-2005 and 2006-2013, and for nonwhites in 2006-2013. Regarding activity and work limitations, and activity and work unable, veterans show statistically higher rates than nonveterans in all survey periods, except the period immediately after the conflict, in 1974-1981. Regarding the current smoker health behavior, white and nonwhite veterans have statistically significant higher incidence in all survey periods (except for nonwhites in 1982-1996) by an average over surveys of roughly 10 pp. As for current drinker (available only in the last two survey periods), veterans also have statistically higher incidence (except for whites in the 2006-2013 survey period) that averages roughly 7 pp over the two survey periods. In sum, the descriptive statistics indicate that veterans are, on average, less healthy and have worse health behaviors relative to nonveterans, but in the case of general health outcomes this is statistically the case only in the medium- to long-run from the end of the Vietnam War. This evidence is in line with studies that report average health outcomes comparisons between veterans and nonveterans (e.g., Kramarow and Pastor, 2012, Gustman et al. 2016).

4 Econometric Methods

Our approach to deal with selection bias and undertake statistical inference on the Vietnam-era military service effects on health outcomes and behaviors uses the Vietnam-era draft lotteries as an IV. We start by offering evidence on the plausibility that the draft lotteries IV does not have a direct effect on the health outcomes, net of its effect working through military service. We call this effect a net average treatment effect (NATE) below, and employ the nonparametric bounds in Flores and Flores-Lagunes (2010, 2013) and Chen et al. (2020) to bound this effect for the subpopulation of never takers. Subsequently, since we find no statistical evidence against the validity of the draft lotteries as an IV, we use standard IV methods to provide point estimates of the military service effects on the health outcomes and behaviors of compliers. Lastly, to conduct statistical inference on the Vietnam-era military service effects for volunteer veterans, we adopt the nonparametric bounding techniques in Chen et al. (2018). These bounds rely on the same assumptions employed by the standard IV estimator (Angrist et al., 1996), plus a relatively mild mean weak monotonicity assumption explained below. Under the same assumptions, the bounding technique allows bounding the average treatment effect on the treated, which in this case corresponds to the entire population of veterans. The exposition of the nonparametric bounding techniques in the balance of this section focuses largely on the intuition behind them. The interested reader is referred to Flores and Flores-Lagunes (2010, 2013) and Chen et al. (2018, 2020) for details.

4.1 Potential Outcomes Framework

Consider a random sample from the population of interest where, for each individual i, we denote the Vietnam-era veteran status by D_i ($D_i = 1$ if i is a veteran and $D_i = 0$ if a non-veteran). Let the exogenously assigned draft-eligibility status be denoted by Z_i ($Z_i = 1$ if eligible, $Z_i = 0$ if ineligible). We denote by D_{1i} and D_{0i} the potential veteran statuses of individual i, where D_{1i} is the potential veteran status if i were to be eligible to be drafted, and D_{0i} if i were not eligible to be drafted.

We partition the total population into four latent principal strata based on the values of the vector $\{D_{1i}, D_{0i}\}$ following Imbens and Angrist (1994) and Angrist et al. (1996). Table 2 illustrates the relationship of these latent strata with the observed groups defined by the observed Vietnam-era draft-eligibility status and the observed veteran status. The first stratum is never takers (nt): individuals who are non-veterans regardless of their eligibility to draft ($D_{1i} = 0$, $D_{0i} = 0$). When never takers receive a high lottery number that was not called, they do not volunteer to serve (upper-left cell in Table 2); and when they receive a low lottery number that was called for induction, they do not serve (upper-right cell in Table 2). For our purposes, there were two relevant reasons why never takers called for induction did not serve.

Given that the military screened prospective recruits for their suitability to military service with, among others, physical, medical, and mental examinations (e.g., Baskir and Strauss, 1978), never takers include individuals whose pre-draft characteristics (e.g., health) prevented them from passing those screening exams. This aspect will be key in justifying one of our assumptions introduced below. In addition, some never takers called for induction could have undertaken strategic actions (e.g., attending college) to avoid the draft. As discussed below, such strategic actions may represent a threat to the exclusion restriction assumption of the IV.

The second stratum is volunteer veterans or always takers (at): individuals who, regardless of whether they are eligible to draft, serve in the military ($D_{1i} = 1$, $D_{0i} = 1$). This is a very important stratum from a policy point of view, as it represented about three-quarters of all servicemen during the Vietnam era. Moreover, the results for this group could also be informative about the U.S.'s current all-volunteer induction system. The third stratum is compliers (c): individuals who serve in the military only if their lottery number is called to enlist ($D_{1i} = 1$, $D_{0i} = 0$). The previous literature using the draft lotteries as an IV to estimate the effects from military service focuses on this stratum, which represents about one-quarter of all Vietnam-era veterans. The last stratum is defiers (df): individuals who enlist when their lottery number is not called for induction, and avoid enlistment if their lottery numbers are called ($D_{1i} = 0$, $D_{0i} = 1$). This stratum will be ruled out below by assumption.

Denote the health outcomes and behaviors by Y_i. For example, $Y_i = 1$ for individuals who have activity limitation and $Y_i = 0$ for individuals without activity limitation. We define all outcomes such that higher values represent worse health (e.g., activity limitation or being a current smoker). Denote by $Y_i(1)$ and $Y_i(0)$ individual i's potential outcomes under treatment $D = d$, that is, the outcome individual i would experience if served in the military or not, respectively. Lastly, let $Y_i(z, d)$ be the potential outcome as a function of the eligibility to draft (the IV) and military service status (the treatment). In what follows, we assume access to data on (Z_i, D_i, Y_i), where $D_i = Z_i D_{1i} + (1-Z_i)D_{0i}$ and $Y_i = D_i Y_i(Z_i, 1) + (1-D_i)Y_i(Z_i, 0)$. Also, to simplify notation, we write the subscript i only when deemed necessary.

4.2 Assumptions

We start with the assumptions used in the standard IV method under heterogeneous effects, which allows identification of the effect of interest for compliers (Imbens and Angrist, 1994 and Angrist et al., 1996). The first assumption, A1, is the random assignment of the instrument Z (the draft-eligibility status): $\{Y(1,1), Y(0,0), Y(0,1), Y(1,0), D_0, D_1\}$ is independent of Z. The Vietnam-era draft lotteries satisfy A1 by design since the lottery numbers were assigned randomly based on birth dates. However, Fienberg (1971) documented an issue related to the randomization mechanism in the 1969 lottery draft: men with birthdays in later months had a higher probability to be drafted (i.e., they tended to receive lower lottery numbers) relative to men with birthdays in earlier months. To account for this issue, the previous literature uses birth month-by-year indicators in estimation (e.g., Dobkin and Shabani, 2009; Angrist et al., 2010). For the same reason, in all estimates to be presented in the next section, we first regress each health outcome on birth month-by-year dummies and survey year dummies, and re-center the corresponding residuals around the health outcome mean. These re-centered residuals are used in all subsequent estimations.

The second assumption, A2, states that there is a non-zero average effect of the draft-eligibility status on veteran status: $E[D_1 - D_0] \neq 0$. A2 is satisfied given the documented positive and statistically significant effect that draft eligibility had on the Vietnam-era veteran status (e.g., Angrist, 1990). The third assumption, A3, is individual-level monotonicity of D in Z: $D_{1i} \geq D_{0i}$ for all i. A3 implies that draft eligibility weakly affects veteran status in one direction, or, equivalently, that there are no defiers (i.e., individuals for whom $D_{0i} = 1, D_{1i} = 0$). A3 is likely satisfied since individuals who prefer enlistment when ineligible-to-draft would also prefer enlistment when they are eligible-to-draft. The fourth assumption, A4, known as the exclusion restriction, states that the draft eligibility IV affects the health outcomes and behaviors exclusively through military service: $Y(0, d) = Y(1, d) = Y(d)$ with $d \in \{0, 1\}$, for all i. This assumption rules out a non-zero net or direct effect of draft eligibility on health outcomes. Given that there are reasons why A4 may not be satisfied in this setting—such as the possibility of individuals undertaking strategic behaviors to avoid serving in the military, if drafted—we will examine whether we can detect statistically significant net average effects of the eligibility to draft on health outcomes and behaviors of never takers, while maintaining A1 to A3, as will be explained below. Having a zero average net effect of the eligibility to draft on health outcomes and behaviors of never takers is a necessary (but not sufficient) condition for the exclusion restriction to hold, as this assumption is imposed at the individual level.[7]

To undertake statistical inference on the effects of military service on the health outcomes and behaviors of volunteers and all veterans, we make use of two additional assumptions, one of which is trivially satisfied. Assumption A5 states that the potential outcomes under consideration are bounded: $Y(d) \in [y^l, y^u]$ for $d = 0, 1$. Given that all of our health outcomes and behaviors are binary, A5 is trivially satisfied as they are bounded between zero and one. The last assumption, A6, is a mean weak monotonicity condition that relates the potential outcomes under no military service of the stratum of volunteers to that of never takers. Specifically, A6 states: $E[Y(0)|nt] \geq E[Y(0)|at]$. Recalling that higher values of the outcomes denote worse health, A6 implies that, in the absence of military service for both strata, the

[7]Note, however, that one can relax this assumption by only requiring it to hold on average within strata instead of at the individual level. One could assume $E[Y(0, d)|k, Z = 0] = E[Y(1, d)|k, Z = 1] = E[Y(d)|k]$ for $d \in \{0, 1\}$ and $k \in \{nt, at, c, df\}$, i.e., the IV does not affect mean potential outcomes within any strata (Huber and Mellace, 2015).

average potential health outcomes of never takers are no better than those of volunteers. We provide a detailed justification of this assumption in section 5.3.2 based on the institutional context and available empirical evidence. For now, we point out that the physical, medical, and mental examinations performed by the U.S. military before enlisting individuals (e.g., Shapiro and Striker, 1970) is a feature that helps to make A6 plausible.

4.3 Identification and Estimation of the Effects of Interest

4.3.1 The Net Effect of the Draft Lotteries on the Health Outcomes of Never Takers

The net (or direct) effect of the draft lotteries on the health outcomes and behaviors of never takers allows us to examine whether the draft lotteries have an average effect on outcomes independently (or net) of their effect working through military service. Given that the exclusion restriction assumption (A4) implies that such effects is zero, finding statistical evidence that such effect is not zero under assumptions A1 to A3 (which are deemed plausible in our application) can be interpreted as a rejection of A4 and, thus, as evidence that the draft lotteries cannot be regarded as a valid IV for military service.[8] Since the net average effect is not point identified under assumptions A1 to A3, our statistical inference on this effect is based on bounds (partial identification). Importantly, because the net effect of the draft lotteries is likely to work through draft avoidance behaviors, we focus on bounding the net effect on the outcomes of never takers. Additionally, partial identification of this net average effect for never takers requires fewer assumptions relative to the same effect for the entire population.

To partially identify the net effect for never takers, we employ the nonparametric bounds developed by Flores and Flores-Lagunes (2010, 2013) and Chen et al. (2020). We start by formally defining the net average effect for never takers following those papers. To do so, consider the counterfactual outcome $Y_i(z, D_{(1-z)})$, which represents the outcome individual i would obtain if he were exposed to the value of the IV of z but his treatment status were under the effect of the IV at the alternate value $(1-z)$. Intuitively, $Y_i(z, D_{(1-z)})$ can be thought as the outcome from an experiment in which the individual is exposed to $Z_i = z$ but the effect of the IV (Z) on the treatment (D) is blocked by holding D fixed at $D_{(1-z)}$. Thus, in our setting, $Y_i(1, D_0)$ represents the counterfactual outcome where the individual is eligible to draft but has the potential veteran status with the value it would have if he was ineligible to draft. Similarly for $Y_i(0, D_1)$. Using these counterfactual outcomes, the net average effect of the draft lotteries on the health outcomes and behaviors of the subpopulation of never takers

[8]For completeness, we also estimated bounds on the net average effects of volunteers (which can also be derived under assumptions A1 to A3). The results for volunteers are similar to the results for never takers presented in section 5.2. These results are available upon request.

is given by:

$$LNATE_{nt}^{z} = E[Y(1, D_z)|nt] - E[Y(0, D_z)|nt], \;\; for \; z = 0, 1,$$

where $LNATE$ stands for "local" net average treatment effect (i.e., local to a subpopulation).[9] The superscript z reflects the fact that this effect is allowed to vary depending on the value of Z, while the conditioning on nt means that it is defined for the latent subpopulation of never takers. The intuition behind this effect is that it measures the difference in potential outcomes when the eligibility-to-draft IV moves from 0 to 1, while the potential value of military service is fixed at a value of the IV of z. Thus, it denotes the causal effect of the IV on the outcome net of its effect working through military service. Lastly, note that for never takers we have $D_0 = D_1 = 0$, which implies that $LNATE_{nt}^{0} = LNATE_{nt}^{1} = LNATE_{nt}$. Therefore, we write the net average effect of never takers simply as:[10]

$$LNATE_{nt} = E[Y(1, 0)|nt] - E[Y(0, 0)|nt]. \tag{1}$$

Partial identification of $LNATE_{nt}$ in (1) relies on assumptions A1 to A3, that is, the same assumptions as the standard IV estimator with the exception of A4 (the exclusion restriction). Flores and Flores-Lagunes (2010, 2013) and Chen et al. (2020) derive bounds on $LNATE_{nt}$ using "trimming bounds" (Zhang and Rubin, 2003; Lee, 2009; Zhang et al., 2008). Intuitively, note that A3 eliminates the defiers stratum, enabling point identification of the first term in (1). Denote $\overline{Y}^{zd} = E[Y|Z = z, D = d]$. Then, the potential outcome of the eligible-to-draft never takers is $E[Y(1, 0)|nt] = \overline{Y}^{10}$). This follows from Table 2 once the d stratum is eliminated. Additionally, the population proportions of the three strata, denoted as π_{at}, π_c and π_{nt}, are also point-identified (Imbens and Angrist, 1994; Angrist et al., 1996): letting $p_{d|z} \equiv Pr(D = d|Z = z)$ for $d, z = 0, 1$, then $\pi_{at} = p_{1|0}$, $\pi_c = (p_{1|1} - p_{1|0})$, and $\pi_{nt} = p_{0|1}$. However, note that we cannot distinguish the never takers from compliers when they are both ineligible-to-draft and did not serve in the Vietnam era (the upper left cell in Table 2). For this reason, the second term in (1), $E[Y(0, 0)|nt]$, is not point identified. Nevertheless, trimming bounds can be constructed on this term. To see this, write the observed average health outcome $\overline{Y}^{00}$ as a function of the average potential health outcomes of the nt and c strata as (Imbens and Rubin, 1997):

$$\overline{Y}^{00} = \frac{\pi_{nt}}{\pi_{nt} + \pi_c} \cdot E[Y(0, 0)|nt] + \frac{\pi_c}{\pi_{nt} + \pi_c} \cdot E[Y(0, 0)|c] \tag{2}$$

[9] Note that, although the literature also refers to the net average treatment effect as the "direct effect", this effect does not have to be direct in any sense—it may still affect the outcomes through channels such as draft avoidance behaviors, as long as these channels are different from the actual military service.

[10] More generally, Flores and Flores-Lagunes (2013) and Chen et al. (2020) decompose the total average effect of the potentially invalid IV (e.g., the eligibility-to-draft) on the outcome into the effect that works through the treatment (military service) and the effect that is net of the treatment. They do this by defining the net average treatment effect ($NATE^{z}$) and the mechanism average treatment effect ($MATE^{z}$) for the entire population and for different strata. In this section, we focus on never takers, whose mechanism average treatment effect is zero because their eligibility to draft does not affect their veteran status (i.e., $D_0 = D_1 = 0$). The reader is referred to those papers for further details on these effects.

Having two unknowns ($E[Y(0,0)|nt]$ and $E[Y(0,0)|c]$), the potential outcome $E[Y(0,0)|nt]$ can be bounded from above by the expected value of the $\frac{\pi_{nt}}{\pi_{nt}+\pi_c} = p_{0|1}/p_{0|0}$ fraction of the largest values of Y in the observed group with $\{Z = 0, D = 0\}$. Similarly, a lower bound on $E[Y(0,0)|nt]$ can be constructed by using the same fraction of smallest values of Y in the same observed group. Since the other component in (1) is point identified, bounds on $LNATE_{nt}$ can be obtained (see Flores and Flores-Lagunes, 2010, 2013 and Chen et al., 2020 for details). Estimation of these bounds relies on plug-in estimators, while statistical inference is based on Imbens and Manski (2004) confidence intervals that cover the parameter of interest with a given probability.[11]

4.3.2 The Local Average Treatment Effect of Military Service on the Health Outcomes of Compliers

Since the conclusion of our assessment of assumption A4 is that we do not find statistical evidence against its implication that $LNATE_{nt}$ in (1) is zero (see section 5.2), we regard the draft lotteries as a valid IV in our analysis of the effect of military service on health outcomes and behaviors of veterans. Under assumptions A1 through A4, the Vietnam-era military service effect on health outcomes and behaviors can be point identified for the latent subpopulation of compliers (Imbens and Angrist, 1994; Angrist et al., 1996). This effect is known as the local average treatment effect or $LATE_c$:

$$LATE_c = E[Y(1)|c] - E[Y(0)|c]. \tag{3}$$

We estimate $LATE_c$ with a standard IV Wald-type estimator, and calculate bootstrapped standard errors. $LATE_c$ is the parameter estimated in prior studies using the draft lotteries IV as source of exogenous variation for military service (e.g., Dobkin and Shabani, 2009; Eisenberg and Rowe, 2009; Angrist et al., 2010). Relative to prior studies, we estimate longer-term effects for compliers of up to 40 years after the end of the conflict when using the last NHIS survey years available, and look at a wider array of health outcomes and behaviors.

4.3.3 The Local Average Treatment Effect of Military Service on the Health Outcomes of Volunteers

While it is possible to point identify the effects of military service on the health outcomes of compliers when using the draft-lotteries IV as exogenous source of variation for military service, compliers represent only about one quarter of the population of Vietnam-era veterans. As can be seen in the bottom row of Table 2, once assumption A3 is imposed (disposing of defiers), the population of Vietnam-era veterans consists of compliers and volunteers. Thus,

[11] The Imbens and Manski (2004) confidence intervals we employ are valid for situations where the width of the bounds on the parameter of interest is bounded away from zero (Stoye, 2009).

volunteers make up the majority of the Vietnam-era veterans—around three-quarters. That compliers represent a relatively small proportion of the population of veterans has been recognized in the literature (e.g., Angrist et al., 1996). Unfortunately, in this context, estimates for the latent subpopulation of volunteers—and in consequence for the population of veterans—are missing in the literature. This is likely because the military service effect for volunteers is not point-identified when using the draft lotteries as an IV, unless additional strong assumptions are employed (such as constant effects). However, the effect can be partially identified under relatively weak assumptions. We take this approach here and employ the nonparametric sharp bounds in Chen et al. (2018), which assume the validity of the draft lotteries IV.

To present the bounds, define the local average treatment effect of military service on the health outcomes and behaviors of volunteers as:[12]

$$LATE_{at} = E[Y(1)|at] - E[Y(0)|at]. \tag{4}$$

It is easy to see why $LATE_{at}$ is not point identified under the conventional IV assumptions A1–A4: while the first term ($E[Y(1)|at]$) is point identified in the data, the second term ($E[Y(0)|at]$) is not. Looking at the lower left cell in Table 2, the first term is identified as $E[Y(1)|at] = \overline{Y}^{01}$. In contrast, the second term in (4) is not point identified simply because volunteers are never observed as nonveterans in the data, making $Y(0)$ for volunteers an entirely hypothetical outcome. Thus, additional assumptions are needed to bound $LATE_{at}$.

Adding assumption A5 in section 4.2, which states that the outcome is bounded, allows to construct "worst-case" bounds (Manski, 1990). These bounds take the following form:

$$\overline{Y}^{01} - y^u \leq LATE_{at} \leq \overline{Y}^{01} - y^l. \tag{5}$$

Unfortunately, these bounds are typically uninformative in practice (e.g, see the empirical application in Chen et al., 2018), and thus it is desirable to entertain additional assumptions. Therefore, we consider nonparametric bounds that impose assumption A6 in section 4.2: $E[Y(0)|nt] \geq E[Y(0)|at]$. We discuss the plausibility of this assumption extensively in section 5.3.2. This assumption improves the upper bound on $E[Y(0)|at]$ by using the corresponding point identified mean for the never takers (nt): $E[Y(0)|nt] = \overline{Y}^{10}$; and given that $E[Y(0)|at]$ enters with a negative sign in (4), the lower bound on $LATE_{at}$ is improved. These bounds take the following simple form:

$$\overline{Y}^{01} - \overline{Y}^{10} \leq LATE_{at} \leq \overline{Y}^{01} - y^l. \tag{6}$$

[12]Note that, in contrast to $LNATE_{nt}$ in (1), where the IV (Z) is allowed to have a net or direct effect on the outcome that does not work through the treatment (D), in the definition of $LATE_c$ and $LATE_{at}$ in (3) and (4), respectively, we make use of the exclusion restriction (A4) by writing $Y(0,d) = Y(1,d) = Y(d)$. In the absence of A4, we would need to specify whether the effect of D on Y is under $z = 1$ or $z = 0$ by writing $LATE_k^z = E[Y(z,1)|k] - E[Y(z,0)|k]$ for $z = 0,1$ and $k = c, at$. For further discussion about the case when A4 is not imposed, see Flores and Flores-Lagunes (2013) and Chen et al. (2020).

Note that relative to the bounds in (5), assumption A6 improves the lower bound only, with the upper bound in (6) being the same as the worst-case upper bound in (5). As a result, our estimated lower bounds on $LATE_{at}$ are expected to be more informative than our estimated upper bounds. This is indeed the case in section 5.3.3 below, and thus most of the discussions about $LATE_{at}$ there will focus on the estimated lower bounds. Estimation of the bounds in (5) and (6) relies on plug-in estimators, and statistical inference on Imbens and Manski (2004) confidence intervals for the parameter of interest.

4.3.4 The Average Treatment Effect of Military Service on the Health Outcomes of Veterans

The population of Vietnam-era veterans is made up of the subpopulations of compliers and volunteers. The military service effect on the health outcomes and behaviors of compliers is point identified, while the effect for volunteers is partially identified. Therefore, with those two elements, we can construct bounds on the military service effect for the population of Vietnam-era veterans, that is, the average treatment effect on the treated (ATT).

Following Angrist (2004) and Chen et al. (2018), the ATT, defined as $ATT \equiv E[Y(1)|D = 1] - E[Y(0)|D = 1]$, can be written as:

$$ATT = \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - E[Y(0)|at])],$$

where $q_1 = Pr(Z = 1)$, $r_1 = Pr(D = 1)$, and $p_{1|0} = Pr(D = 1|Z = 0)$. Noting that the only non-point identified term is $E[Y(0)|at]$, the same term as in (4), it is easy to see how assumptions A5 and A6 provide bounds on the ATT. Under A1 to A5, the worst-case bounds are:

$$\frac{1}{r_1}[q_1(E[Y|Z=1]-E[Y|Z=0])+p_{1|0}(\overline{Y}^{01}-y^u)] \leq ATT \leq \frac{1}{r_1}[q_1(E[Y|Z=1]-E[Y|Z=0])+p_{1|0}(\overline{Y}^{01}-y^l)], \tag{7}$$

while under A1 to A6 the bounds take the following form:

$$\frac{1}{r_1}[q_1(E[Y|Z=1]-E[Y|Z=0])+p_{1|0}(\overline{Y}^{01}-\overline{Y}^{10})] \leq ATT \leq \frac{1}{r_1}[q_1(E[Y|Z=1]-E[Y|Z=0])+p_{1|0}(\overline{Y}^{01}-y^l)] \tag{8}$$

As with $LATE_{at}$, note that A6 improves only the lower bound relative to the worst-case bounds in (7) and, therefore, our estimated lower bounds on the ATT are expected to be more informative that the estimated upper bounds. Finally, as before, the previous bounds are estimated with plug-in estimators and statistical inference relies on Imbens and Manski (2004) confidence intervals.

5 Results

In this section, we present and discuss ordinary least squares (OLS), intention-to-treat (ITT), and the estimates of the four parameters of interest described in the previous section. At the outset, we briefly discuss the estimated stratum proportions under assumptions A1 to A3, which are shown in Online Appendix B for whites and nonwhites born between 1948 and 1952, by each of the survey periods. The largest estimated stratum proportion is that of never takers, accounting for about 64 to 66% of the population across the four survey periods (1974-1981, 1982-1996, 1997-2005, and 2006-2013). The estimated proportion of compliers is about 13 to 15% of the population across survey periods, with the estimated proportion of compliers being considerably lower among nonwhites: about 7 to 9% across survey periods.[13] The group of compliers is the subpopulation for which standard IV estimates point-identify the effect of military service. Lastly, the estimated proportion of volunteers (always takers) is about 21 to 23% of the population across survey periods, with the estimated proportion of volunteers among nonwhite being between 17 to 21%. Given that the veteran proportion is between 26.7 to 28.7%, volunteers represent between 76.4 and 80.2% of the population of veterans across survey periods for the combined group of whites and nonwhites.

Since we analyze the effects of interest on multiple health outcomes, we perform multiple tests of the null hypothesis that each of the effects on the health outcomes equals zero. It is well known that, without appropriate adjustment, this situation increases the risks of falsely rejecting some of these null hypotheses. To avoid this issue, we adopt a conservative adjustment for sequential multiple testing in the sequential False Discovery Rate (FDR) by Benjamini and Hochberg (1995).[14] To implement this multiple testing procedure to the estimated bounds, we obtain the p-value at which each null hypothesis of a zero effect is rejected by the Imbens and Manski (2004) confidence intervals, and then implement the FDR procedure on the total null hypotheses tested across the health outcomes within a given family. The families of outcomes we consider correspond to the four groups of outcomes described in Section 3: general health outcomes, health behaviors, activity limiting chronic conditions, and other chronic conditions. For brevity, in the results described below we only point out instances where the multiple testing procedure overturns the conclusion from the standard inference for the corresponding estimates. Online Appendix C contains the details of the multiple testing procedure and a summary of its results.

[13]This difference in estimated complier proportion has been documented in prior literature, e.g., Dobkin and Shabani (2009) and Angrist et al. (2010).

[14]The FDR measures the probability that all the rejections of a "family" of null hypotheses are false.

5.1 Ordinary Least Squares and Intention-to-Treat Estimates

Table 3 presents OLS and ITT estimates on the groups of general health outcomes and health behaviors for the different survey years under analysis, with the sample broken down by whites and nonwhites.[15] The OLS and ITT estimates for the other two groups of health outcomes we analyze (activity-limiting chronic conditions and other chronic conditions) are shown in the Online Appendix A. The OLS estimates represent the difference in average outcomes between veterans and nonveterans that do not control for the selection problem, while the ITT estimates represent the reduced-form average effect of the draft eligibility status IV on the corresponding outcomes.

The OLS results in Table 3 show that, for both whites and nonwhites, veteran status tends to be statistically significantly associated with worse average health outcomes (i.e., positive estimates) in survey periods after the first one (1974-1981), which span 9 to 40 years after the end of the Vietnam War. To have an idea of the magnitude of these OLS estimates, consider the outcome activity limitation for whites in the latest survey of 2006-2013. The estimate indicates that the detrimental association between veteran status and activity limitation is 4.1 pp, which represents 35.7% relative to the mean activity limitation of nonveterans in Table 1. In contrast, few of the ITT estimates of the reduced-form effect of the draft-eligibility status are statistically significant, mostly for whites, and four of those denote improvements in health (all in the 1997-2005 survey). While finding negative ITT estimates may be somewhat unexpected, in principle, it is possible that draft-eligibility status improves general health conditions in the 1997-2005 survey—for example, if military service improved access to health care, self-discipline, and physical fitness.[16] However, the lack of a clear pattern across periods in the ITT estimates is notable: those statistically negative ITT estimates become statistically insignificant in the 2006-2013 period, and several are positive in periods prior to 1997-2005, with the ITT estimates for activity limitation being statistically positive. As a result, we conclude that the ITT estimates do not provide consistent evidence of statistically significant effects of draft eligibility on general health outcomes and behaviors. In what follows, we aim to go towards causal inference on the effect of military service on health conditions of compliers and volunteers by employing the draft-eligibility status as an IV.

[15] Recall that some outcomes are not available in every survey period analyzed due to changes in the interview instrument, as discussed in section 3.

[16] Dobkin and Shabani (2009) also obtain ITT effects that are negative in their analysis of the 1997-2004 survey period.

5.2 The Net Effect of the Draft Lotteries on the Health Outcomes and Behaviors of Never Takers

In this section, we start by assessing whether draft eligibility has a net or direct effect on the outcomes we analyze that does not work through military service. Figure 1 presents the estimated bounds on the net average effect of the draft lotteries on the general health outcomes and health behaviors of the never takers ($LNATE_{nt}$) under assumptions A1 to A3, which were described in section 4.3.1.[17] As discussed there, a statistically significant non-zero net average effect for this subpopulation implies that assumption A4 (the exclusion restriction) is violated. Each of the panels in Figure 1 corresponds to a single general health outcome or health behavior for either whites (left panels) or nonwhites (right panels). Within each panel, estimated bounds (represented by the shaded bars) and their 95% confidence intervals (represented by vertical lines) are shown for each of the survey periods in which the general health outcome or behavior is available. As can be seen throughout the panels in Figure 1, none of the estimated bounds exclude a zero net effect of the draft lotteries (except for smoking in the 1982-1996 survey period in Panel L), even though in some instances the estimated upper bounds are in principle consistent with large net effects. Correspondingly, none of the 95% confidence intervals exclude zero. Estimated bounds for the activity-limiting chronic conditions and the other chronic conditions outcome groups are shown in Online Appendix A, Tables A.10-A.12. The message from those estimated bounds is also that the net effect is not statistically different from zero according to the 95% confidence intervals. In sum, we are not able to reject that the net average effect of the draft lotteries is zero for the outcomes we analyze. Taken together, we regard this evidence as suggesting that the exclusion restriction (assumption A4) is plausible in the current context. This interpretation of the evidence agrees and complements similar findings in Kitagawa (2015) and Mourifie and Wan (2017), both of which do not find evidence against implications of the combination of assumptions A1 to A4 in the context of analyzing military-service effects on labor earnings outcomes using the Vietnam-era draft lotteries as an IV for military service. Still, we note that, similar to Kitagawa (2015) and Mourifie and Wan (2017), our assessment of assumption A4 only tests an implication of the assumption under assumptions A1 to A3, and not the assumption itself, as A4 is not directly testable.[18]

[17] For all the figures presented in the paper, the corresponding estimates can be found in Online Appendix D.

[18] We have also estimated bounds on the effects of military service on the health outcomes of compliers and volunteers while not assuming A4 (i.e., allowing the IV to violate the exclusion restriction), based on the methodology in Chen et al. (2020). While the bounds disposing of the exclusion restriction are predictably wider and less precise, the main general conclusions we report below in sections 5.3.3 and 5.3.4 hold using those bounds.

5.3 The Effects of Military Service on Health Outcomes and Health Behaviors of Veterans

In this section, we provide our results related to the military service effect on the health outcomes and health behaviors of veterans. We start in section 5.3.1 with the military service effects for compliers, which are point identified under assumptions A1 to A4. Before presenting the novel results for the group of volunteers in section 5.3.3, in section 5.3.2 we discuss the plausibility of the key assumption we employ to bound these effects. To close, we discuss in section 5.3.4 the findings for the entire group of veterans, which combines the subpopulations of compliers and volunteers.

5.3.1 The Effects of Military Service on Health Outcomes and Health Behaviors of Compliers

The military service effects on the health outcomes and health behaviors of compliers are estimated using standard IV regression. Figure 2 presents the IV estimates of the military service effects for compliers on the outcome groups of general health outcomes and health behaviors. As in Figure 1, each panel in Figure 2 corresponds to a single general health outcome or health behavior for either whites or nonwhites. Within each panel, point estimates (represented by a cross), 95% confidence intervals (represented by vertical lines) and the control mean $E[Y|D=0]$ (represented by a bold dot) are shown for each of the survey periods in which the outcome is available. Inspecting the various panels in Figure 2, it is evident the absence of a clear pattern for the effects of military service on the general health outcomes and health behaviors of compliers. The majority of the estimated effects are statistically insignificant—the vertical lines representing the confidence intervals cross the horizontal line representing zero. Meanwhile, the point estimates are sometimes positive (indicating health worsening) and others negative (indicating health improvements).

Some of the few statistically significant effects are as follows. In Panel A, for white compliers, activity limitation is statistically significant during the 1974-1981 and—marginally so—during the 1982-1996 survey periods (the latter effect is no longer statistically significant after the adjustment for multiple testing). These estimates indicate an increase of 6.6 pp (45.8% of the rate of nonveterans) in the first survey and of 4.2 pp (35%) in the second. However, the estimate of the same effect in the 1997-2005 survey period (-7.4 pp; 67.7%) changes sign and is also statistically significant. Other effects for compliers that are statistically significant suggest that military service has a health-improving effect. These estimated effects are on the incidence of fair/poor health in Panel E (-9.3 pp; 106%); the incidence of work limitation in Panel G (-8 pp; 80%); and the incidence of work unable in Panel I (-8 pp; 130%). For nonwhite compliers, none of the estimated effects on general health outcomes is statistically significant in any of the survey periods. While it is possible that military service can have health-improving effects (e.g., through access to above-average health care after service), the lack of a clear pattern over time is noteworthy. All four statistically significant health-improving effects occur for white compliers in the 1997-2005 survey period, but all become statistically insignificant in the 2006-2013 survey period.[19] Panels K through N in Figure 2 present the estimated military service effects on the health behaviors of compliers, where none of the estimates is statistically significant.[20]

To summarize, the point estimates of the Vietnam-era military service effect on the health of complier veterans do not provide consistent evidence of statistically significant effects. These results are in line with prior literature that largely finds weak or no health effects of Vietnam-era military service using the same strategy. For instance, Angrist et al. (2010), using the 2000 Census one-in-six file, found small military service effects on disability rates for whites (about 1.2 pp), and no statistically significant effects for nonwhites. And while they found military service effects on disabilities related to vision or hearing-related problems (between 1.1 and 3.9 pp), they reported no effects on disabilities related to mental health, physical or mobility reasons. Dobkin and Shabani (2009), using the NHIS survey 1974-2004, found that, for most of the health outcomes they consider, the military service effects on compliers is statistically insignificant and imprecisely estimated. Eisenberg and Rowe (2009), using also NHIS survey data, find that military service has a statistically significant effect on smoking in the 1978-1980 survey period of 35 pp, but that these effects become small and statistically insignificant in the long-run (1997-2005).

5.3.2 Discussion and Indirect Assessment of Assumption A6: Ranking of Potential Health Outcomes Between Never Takers and Volunteers

Before presenting results for the military service effect on the health of volunteers and veterans, we discuss and assess in detail the key assumption employed to partially identify their effect, and present some indirect statistical evidence to gauge its plausibility. Assumption A6 states that, in the absence of military service, the average potential health outcomes of never takers are no better than the average potential health outcomes of volunteers: $E[Y(0)|nt] \geq E[Y(0)|at]$. It is important to stress that this comparison of average potential health outcomes between never takers and volunteers is required to hold only under the absence of military service.

[19] Dobkin and Shabani (2009) also documented health-improving military service effects for compliers in the 1997-2004 survey period, but theirs were statistically insignificant. The lack of precision is likely because of a combination of them having a slightly shorter survey period (ours is 1997-2005), a smaller set of cohorts (1950-1952 versus 1948-1952) and combining whites and nonwhites.

[20] Corresponding estimates for activity-limiting chronic conditions and other chronic conditions, presented in Online Appendix A, Tables A.13-A.15, yield similar inconsistent evidence about the military service effects of compliers.

Discussion of Assumption A6

Conventional models of selection into the military, such as human capital models (e.g., Fisher, 1969) and the Roy model (Roy, 1951), typically consider lifetime earnings as the outcome of interest. These models can be adapted to health outcomes (e.g., Grossman, 1972). One implication of these models is that individuals choose the military because they expect to have better health outcomes in that sector. These models do not offer a sharp implication either way about Assumption A6, regardless of whether one builds into them the medical and physical examinations that individuals must pass to enlist.[21] Thus, human capital and sorting models are not inconsistent with Assumption A6.

If one considers the medical and physical screening examinations performed by the U.S. military before enlisting individuals, and their documented stringency (e.g., Shapiro and Striker, 1970; Suttler, 1970), the argument in favor of the plausibility of Assumption A6 is substantially strengthened.[22] There is ample evidence that the failure rate of the medical and physical screening examinations was considerably high. Official Selective Service System records (Selective Service System, 1970, 1971a, 1971b, 1972a, 1972b, 1973) indicate that the pre-induction examinations failure rate ranged between 47% and 57%, while the induction examinations (applied to those who passed the pre-induction examinations) ranged between 21% and 36%.[23] To see how these examinations relate to Assumption A6, first note that, by implication, all volunteers would have taken and passed the medical and physical examinations. Thus, the health of *all* volunteers (and, trivially, volunteer's average health) at the time of the draft is above the threshold implied by these examinations.

The analysis of the average health of never takers under the medical and physical screening examinations is more nuanced, given the endogenous choice to take the examinations. This choice is likely affected by eligibility to draft. To take into consideration this endogenous choice, we employ a principal stratification approach to simultaneously model the endogeneity of both veteran status and the decision to take the examinations.[24] To explain in brief (Online

[21]The lack of a sharp implication of these models toward Assumption A6 follows from the fact that, in them, selection into the military sector by a given individual is based on the perceived effect of military service on health (i.e., $Y_i(1) - Y_i(0)$), whereas Assumption A6 only involves the comparison of health under no military service (i.e., $Y_i(0)$) between two subpopulations. A formal illustration of this can be found in Online Appendix E.

[22]Reinforcing the implications of the failure of the medical and physical screening examinations is the failure of mental aptitude screening tests (e.g., the Armed Forces Qualification Test). Individuals failing these aptitude tests, which correlate with innate ability, would likely attain less education on average. Since education is positively associated with future health (e.g., Zajacova, 2006; Oreopoulos and Salvanes, 2011), men failing these mental aptitude tests, whom become never takers, would have, on average, worse adult health.

[23]These figures agree with reports elsewhere, such as in Angrist (1990, 1991) who reports that half of all registrants in 1970 failed the pre-induction examinations, and 20% of those who passed were eliminated by physical inspections conducted at induction; and Baskir and Strauss (1978) who report that, from 1967 to 1973, the failure rate for the pre-induction physical exam was 47%.

[24]The principal stratification approach to account for more than two endogenous variables is in the spirit of

Appendix F provides the details), let S denote whether the individual decides to take the examinations, and let $S(z)$ be the associated potential values as a function of eligibility to draft Z. After assuming monotonicty of S in Z (which is analogous to Assumption A3 and similarly justified), there are three principal strata with respect to S: NN (respectively, SS) comprises individuals who would never (always) take the examinations regardless of draft eligibility; and NS comprises individuals who would take the examinations only if drafted. Stratifying the population in terms of both D and S results in the following strata: $\{nt, at, c\} \times \{NN, SS, NS\}$, where, for example, the stratum $ntNN$ comprises never takers who, regardless of draft eligibility, decide to not take the examinations. Comparisons of individuals within strata are free of endogeneity biases, just as with the compliance types in IV analysis (Angrist et al., 1996).

To analyze the average health of never takers the three relevant strata are: (1) never takers who would take the examinations regardless of eligibility to draft ($ntSS$), (2) never takers who would take the examinations only if they are eligible to draft ($ntNS$), and (3) never takers who do not take the examinations regardless of eligibility to draft ($ntNN$).[25] The average health of never takers will be a weighted average of these three strata. After discussing the average health of each of the strata at the time of the draft lotteries (and their relationship to future health), we will quantify the proportion of never takers in each of the three strata. This will help us shed light on the average health of never takers and thus on the plausibility of Assumption A6. The additional assumptions imposed in this analysis are (1) the already mentioned monotonicity of the choice to take the examinations (S) in the eligibility to draft (Z), and (2) that never takers who obtain deferments or engage in draft-avoidance behaviors do not do so *after* having passed the examinations.[26]

Consider the first stratum of never takers, $ntSS$. Since men in this group take the examinations regardless of eligibility to draft, they can be thought of as "patriots" who would be

Chen and Flores (2015) and Blanco et al. (2020). In short, it consists of stratifying individuals into latent groups based on both veteran status and whether the individual takes the medical and physical examinations. The approach, proposed by Frangakis and Rubin (2002), is rooted in the LATE framework (Imbens and Angrist, 1994; Angrist et al., 1996).

[25] Note that, in this context, the principal stratification $\{nt, at, c\} \times \{NN, SS, NS\}$ reduces to the following five strata: $ntSS$, $ntNS$, $ntNN$, $atSS$ and cNS. The reason there is only one stratum of volunteers is that volunteers serve—and thus take and pass the medical examination—regardless of eligibility to draft, so the strata $atNN$ and $atNS$ are ruled out. Similarly for compliers, they serve only if drafted, implying that the only non-empty stratum is cNS. Thus, one can refer to the strata $atSS$ and cNS simply as at and c, respectively.

[26] Both of these assumptions are likely mild. The monotonicity assumption is justified along the same lines as the justification of Assumption A3. The second assumption is natural since deferments and draft avoidance would likely happen before engaging with the military by taking the medical and physical examinations. Indeed, the steps in the conscription process (see, e.g., Shapiro and Striker, 1970; 32 C.F.R. §§1622.2, 1623.2) provide for opportunities to obtain deferments before undergoing the examinations. Moreover, for men undertaking draft avoidance actions, it does not seem desirable to present themselves for pre-induction and induction, take the examinations, and subsequently engage in draft avoidance actions.

volunteers had there been no examinations. It is their failing of the examinations that make them never takers. Importantly, given that all men in this stratum take the examinations and fail them, their individual health will be below the threshold implied by these examinations. Consequently, the $ntSS$ stratum will certainly have lower average health at the time of the draft relative to the volunteers.

The second stratum of never takers, $ntNS$, comprises individuals who would take the examinations only if eligible to draft. Note that the group of individuals in this stratum who actually take (and thus fail) the test is chosen at random because $Z = 1$ is randomly assigned. However, by definition of the stratum, all the $ntNS$ individuals would have taken and failed the examinations if they had been drafted, since they take the examinations and do not serve under $Z = 1$. As a result, the average health of this stratum will be below the threshold implied by the examinations. Consequently, the average health of this stratum at the time of the draft is also below that of the volunteers.[27]

The analysis above suggests that the previous two strata have worse average health relative to volunteers at the time of the draft. To link health at the time of the draft with the potential outcomes under no military service for never takers, $Y(0)$, we note that there is a strong association between early-life health and health later in life (e.g., Banks et al., 2012). Thus, we would expect the average potential health outcomes of volunteers, had they not served, to be no worse than those of the individuals failing the medical, physical, and mental examinations employed by the military before induction.

Turning to the third stratum of never takers, the $ntNN$ comprises individuals who decide not to take the examinations regardless of their eligibility to draft. Given the random probability of being eligible to draft, 42.4% of men in $ntNN$ are expected to be eligible to draft and thus they must have chosen to either obtain an allowable deferment or undertake draft avoidance behaviors to bypass taking the examinations. Conversely, 57.6% of them did not have to obtain an allowable deferment or undertake draft avoidance behaviors (although some of them might have done so) since they did not face conscription. The distinction is relevant since some of these actions may be conducive to worse future health outcomes, while others may be conducive to better future health outcomes. One relevant draft-avoidance action was to engage in delinquencies or criminal activities, since having a criminal record was a way to avoid being drafted into military service by failing the military induction's "moral standards"

[27] The discussion in this paragraph implicitly assumes that there are no individuals in the $ntNS$ stratum who, if not drafted, would be healthy enough to pass the medical and physical screening examinations. We can think of two reasons why this may not be the case. The first is due to individuals who would hurt themselves to fail the examinations if they are drafted. The second involves convincing or bribing doctors to certify a bogus medical condition. While there is anecdotal evidence that those two practices existed (e.g., Baskir and Strauss, 1978), it is hard to imagine that they were so widespread as to reverse the conclusion that the average health of the $ntNS$ stratum is lower than that of volunteers at the time of the draft, and more generally, to invalidate assumption A6 given the discussion and empirical evidence presented in this section.

(Suttler, 1970; Shapiro and Striker, 1970). Also, being drafted and simply refusing to report for induction lead to convictions and prison sentences for draft offenders following the draft law (Baskir and Strauss, 1978).[28] We would expect never takers who engaged in these draft-avoidance behaviors to have worse average potential health, given existing evidence that early incarcerations correlate with worse future health outcomes (e.g., Massoglia and Pridemore, 2015).

Conversely, other actions, such as the educational and paternity deferments, could in principle be related to better future health outcomes. Regarding educational deferments, which could lead to better adult health given the positive correlation between schooling and health (e.g., Zajacova, 2006; Oreopoulos and Salvanes, 2011), there are several reasons to expect that they will not dramatically drive up the average health of eligible-to-draft never takers. First, from the available evidence, the group of individuals who took educational deferments is unlikely to represent a large proportion of eligible-to-draft never takers, particularly relative to those never takers who failed medical and physical examinations, mental aptitude screening tests, failed the military induction's moral standards, or just avoided the draft without justification.[29] Second, the empirical evidence does not support the notion that the effect of the draft lotteries on educational attainment is large enough to statistically significantly affect the health and labor market outcomes of never takers. Specifically, if the educational deferment resulted in significant health benefits for never takers as a group, we would expect to find empirical evidence that the net average treatment effect of the draft lotteries on their health and labor market outcomes is different from zero (and thus that the exclusion restriction is violated). As discussed in section 5.2, the data is not able to reject the null hypothesis that such effect is zero for the health outcomes and behaviors we consider, while other work (Kitagawa, 2015; Mourifie and Wang, 2017) also failed to find evidence that the exclusion restriction does not hold when looking at labor earnings outcomes. Regarding the paternity deferments, Kutinova (2009) documents that this deferment was frequently used until April of 1970, when it was eliminated. Still, official statistics (Selective Service System, 1970, 1971a, 1971b, 1972a, 1972b, 1973) indicate that a non-negligible proportion of deferred individuals fell in this category in the draft lottery years of 1970, 1971, and 1972. Given that

[28] Peterson (1998) documents that almost half of the 570,000 traceable individuals who were eligible to draft and did not present themselves for induction during the U.S. Vietnam War became accused draft offenders, and around 22,000 of them were convicted after being brought to trial.

[29] For instance, Card and Lemieux (2001) found that draft avoidance increased college attendance by between 4 and 6 pp, the fraction of individuals with some college by about 4 pp, and the fraction with a college degree by about 2 pp. While statistically significant, the magnitude of these effects is unlikely to yield never takers who took educational deferrals the main group among the $ntNN$ never takers. Moreover, the arguments in Card and Lemieux (2002) about the effects of the draft process on educational deferments imply that their estimates are likely driven by men exposed to the draft process *prior* to the introduction of the draft lotteries (which applied to all men aged 19 to 25) and not by men exposed to the draft lotteries (which was limited to men in particular year cohorts). Further supporting this view, Angrist and Chen (2011) argue that educational deferments were probably of little importance for the Vietnam-era draft cohorts.

this particular deferral mechanism ended in April 1970 and that there is a documented sharp decrease in such deferments from 1970 to 1972 (the years of inductions in our sample), it seems plausible that a large proportion of paternity deferments correspond to earlier cohorts rather than to the cohorts we consider. Moreover, in contrast to the strong positive correlation of education and future health, in the case of paternity the literature is mixed about the sign of the correlation of paternity with future health (e.g., Torche and Rauf, 2021).[30] In sum, given the differential impact on adult health of the various kinds of deferments and draft avoidance behaviors undertaken by a fraction (approximately 42.4%) of men in the $ntNN$ stratum, it is difficult to conjecture about the average potential health outcomes of this stratum relative to that of volunteers.[31]

The average health of never takers as a whole is a weighted average of the average health of the three strata previously discussed, with weights given by the proportion that each stratum represents of the group of never takers. Denote these weights as $\delta_k = \pi_k/\pi_{nt}$, where π_k denotes stratum k's proportion in the population for $k = \{ntSS, ntNS, ntNN\}$. It is not possible to estimate these three weights (δ_k) without imposing controversial assumptions; however, they can be bounded under the current assumptions using available data. The details on how the weights can be bounded are shown in Online Appendix F. Using information on the probability of failing the examinations conditional on taking them (58%),[32] the fraction of never takers who failed the examinations (72.1%),[33] and the steps outlined in Online Appendix F, we can derive the following bounds on the three weights: $\delta_{ntSS} \in [50.5\%, 72.1\%]$, $\delta_{ntNS} \in [0, 49.5\%]$, and $\delta_{ntNN} \in [0, 27.8\%]$.[34]

Based on these bounds on the weights, the stratum $ntSS$, for which it is certain that its average health is considerably lower than that of volunteers at the time of the draft, has the

[30]Toche and Rauf (2021) use NLSY data with models that include individual fixed effects to account for possible selection. Their results suggest that paternity is related to less tobacco, alcohol, and drug use, but also to higher weight and lower self-reported health.

[31]We note that the health of some of the individuals in the $ntNN$ stratum may be such that if they were to take the examinations they would fail them. They could have decided not to take the examinations for fear of passing them, and thus obtained a deferment or engaged in draft avoidance behaviors. Therefore, it is not necessarily the case that all individuals in the $ntNN$ stratum are above the examinations' health thresholds.

[32]This is a *conservative* estimate of the overall failure rate, computed from the conservative 47% pre-induction failure rate and the 21% induction failure rate of the medical and physical examinations obtained from the Selective Service System statistics previously cited. Thus, $0.47 + [(1 - 0.47) * 0.21] = 0.5813$.

[33]This is computed using the NHIS 1974-1981 (see Online Appendix Table B.1 and Online Appendix F) by (1) obtaining the total number of people who took the examinations by dividing the total number of veterans (10,374) by the examination's passing rate (42%) which yields 24,700; (2) obtaining the total number of people who failed the test as $24,700 - 10,374 = 14,326$, whom are never takers that decided to take the examinations; and (3) dividing (2) by the total number of never takers: ($\frac{14{,}326}{19{,}862} = 0.721$)—the total number of never takers is obtained by multiplying the population ($33,103$) by the estimated proportion of never takers in the population ($\pi_{nt} = 0.60$.)

[34]Bounds on the corresponding stratum proportions (π_k) can be obtained by multiplying these bounds by $\pi_{nt} = 0.60$.

highest lower bound of 50.5% and an upper bound of 72.1%. The stratum $ntNS$, which is very likely to have lower average health than volunteers at the time of the draft, has a lower bound of zero and an upper bound of 49.5%. Of particular interest is the upper bound of the $ntNN$ stratum, given that there is more uncertainty about its average health relative to that of volunteers. This stratum represents, at most, 27.8% of the never takers. Moreover, it seems unlikely that this upper bound is reached, as this would require that $\delta_{ntNS} = 0$ (see the derivations in Online Appendix F), which would rule out the existence of never takers who would take the examinations only if drafted. Having $\delta_{ntNS} = 0$ would also imply that all individuals who would present for induction and take the examinations only if drafted would pass them, which seems unlikely.

Summarizing the discussion of Assumption A6, our analysis of never takers classifies them into three strata to account for the possible endogenous choice of taking the medical and physical examinations. The average pre-draft health of the group of never takers is a weighted average of the average health of the three strata, with weights given by their proportion in the group of never takers. Our analysis indicates that, relative to volunteers, one stratum ($ntSS$) is certain to have lower average health, another is very likely to have lower average health ($ntNS$), while the last ($ntNN$) is somewhat more uncertain to tell. We are able to bound the proportion of each stratum, indicating that the stratum with more uncertain average health represents a relatively small fraction of the group of never takers. As a result, the average health of the $ntNN$ stratum would have to be dramatically better than that of the volunteers to push the overall average health of never takers above that of volunteers, which seems very unlikely. Therefore, it is very likely that the group of never takers as a whole has lower average health relative to the group of volunteers at the time of the draft.[35] Given the strong association between early-life health and health later in life (e.g., Banks et al., 2012), we would expect the average potential health outcomes of volunteers, had they not served, to be no worse than that of never takers, lending support to assumption A6. Consequently, we regard assumption A6 as plausible given the available data and the institutional context.

Indirect Statistical Evidence on Assumption A6

We now present indirect empirical evidence that supports assumption A6. Note that, since $E[Y(0)|at]$ is not observable in the data, it is not possible to directly test the validity of assumption A6. However, indirect evidence can be brought to bear by estimating the average of selected characteristics for each of the latent subgroups prior to the draft lotteries. These averages can be estimated under assumptions A1 to A3. If those pre-draft characteristics are

[35]This conclusion is consistent with the extant literature arguing that individuals who self-select into the military are "positively selected" in the sense that they are healthier on average than individuals who do not self-select into the military (e.g., Seltzer and Jablon, 1974; Bedard and Deschênes, 2006; Eisenberg and Rowe, 2009).

highly correlated to health outcomes and their averages are consistent with the condition in assumption A6, then they represent indirect evidence consistent with the assumption. This exercise is in the spirit of Flores and Flores-Lagunes (2010, 2013) and Chen et al. (2018, 2020), among others, who inform similar mean weak monotonicity assumptions in other contexts. We concentrate on two relevant pre-draft characteristics that are available in our data. The first one is incidence of activity limitations before 1965. Given that healthier children tend to become healthier adults (e.g., Almond and Currie 2010; Banks et al., 2012), it is reasonable to assume that a stratum that has a higher rate of activity limitations before the draft lotteries (i.e., while aged 12-18 years old) would have no better mean potential health outcomes compared to a stratum with a lower pre-draft rate of activity limitations, when both are compared under no military service. The second pre-draft characteristic is high school incompletion, which is relevant since prior studies document a positive correlation between schooling and health (e.g., Zajacova, 2006; Oreopoulos and Salvanes, 2011). Therefore, a stratum that has higher pre-draft high school incompletion rates is expected to have no better mean potential health outcomes under no military service compared with a stratum that has lower pre-draft high school incompletion rates.

The variable "activity limitations before 1965" is constructed by using two available variables in the 1974-1981 survey period: "limitation of activity" and "duration of limitation". We assign the value of one to the variable activity limitations before 1965 if the duration of the limitation is larger than the corresponding survey year response subtracted by 1965, and assign the value of zero otherwise (including if the respondent is not limited in any way). Unfortunately, we cannot construct this variable for the 1982-1996 survey period because the duration of limitation variable is not available in this survey. The variable high school incompletion is constructed based on individuals who have not completed 12 years of education and above. We focus on the earlier 1974-1981 and 1982-1996 survey periods in an attempt to minimize the impact of the GI Bill on education (e.g., Card and Lemieux, 2001; Angrist and Chen, 2011). We do this even though there is evidence in Angrist and Chen (2011) that Vietnam-era military service increased primarily the years of college attendance among compliers, having comparatively smaller impact on high school completion.

Table 4 presents the two estimated average pre-draft characteristics described above. The first three columns correspond to the three different latent subpopulations (never takers, volunteers, and compliers). The last three columns show the three corresponding differences in the averages between the subpopulations. Even though assumption A6 involves only the subpopulations of never takers and volunteers, compliers are included for completeness. The four horizontal panels correspond to the combinations of survey period (1974-1981 or 1982-1996) and white/nonwhite demographic group.

The indirect assessment of assumption A6 using average pre-draft characteristics in Table 4 is suggestive of the plausibility of the assumption. This can be seen in the fourth column in Table 4, which shows the difference in average pre-draft characteristics between volunteers and never takers. All of the differences are negative and highly statistically significant, indicating that never takers have higher activity limitations before 1965 and higher rates of high school incompletion, relative to volunteers. Moreover, most of those differences are large in magnitude. Given the natural positive correlation of adolescent health and adult health, and the positive correlation between schooling and health, these results are consistent with assumption A6.

Looking at the other two differences in average pre-draft characteristics reported in Table 4 (columns five and six), it is interesting to note that never takers also have higher rates of activity limitation before 1965 and high school incompletion relative to the group of compliers. All the estimated differences are positive and statistically significant (with the exception of activity limitation before 1965 for nonwhites in the 1974-1981 survey period, which is positive but not statistically significant). This evidence is consistent with the previous arguments that never takers are likely the stratum with the worse average health, and that this is likely influenced by the physical, medical, and mental aptitude screening tests that take place prior to induction. Lastly, the difference in average pre-draft characteristics between volunteers and compliers suggest that volunteers and compliers have very similar rates of activity limitation before 1965, as the difference between the two subgroups is essentially zero. Meanwhile, volunteers have higher rates of high school incompletion relative to compliers, which are statistically significant among whites.

In sum, based on the arguments about the relative health of never takers, plus the indirect evidence from average pre-draft characteristics, we conclude that it is reasonable to assume that never takers would have no better average potential long-term health outcomes relative to volunteer veterans, had they not served. The evidence gives us confidence that assumption A6, which is key to derive informative lower bounds on volunteers and the population of veterans, is plausible and likely a relatively mild assumption.

5.3.3 The Effects of Military Service on Health Outcomes and Behaviors of Volunteers

After having discussed and assessed the plausibility of the key assumption A6, we present the estimated bounds for volunteers. Since these results are novel in this literature, we present them for all health outcomes and behaviors available in the NHIS. Figure 3 shows the estimated bounds on the effects of military service on the general health outcomes and health behaviors of volunteers. The panels in Figure 3 mirror the format of previous figures, separately presenting results for whites and nonwhites, with each panel containing the estimated bounds for each survey period in which a corresponding outcome is available. The general pattern of statistical inference on the military service effect of volunteers is as follows. In the few years following the end of the Vietnam War, the estimated bounds (and their confidence intervals) include zero, except for current smoker. However, for the subsequent survey periods the estimated bounds clearly imply important detrimental military service effects on the general health outcomes and health behaviors of volunteers, with the statistical significance of the effects generally increasing over time. Given that these effects can only be partially identified under our assumptions, even though the statistical significance of the effects and the estimated lower bounds generally increase over time, it is not possible to determine whether the magnitude of the effects is statistically increasing over time.

We illustrate this general pattern with the military service effect of volunteers in the context of a selection of health outcomes and behaviors from Figure 3.[36] Consider Panels A and B in Figure 3, which show the estimated bounds on activity limitation for whites and nonwhites, respectively. In Panel A, the estimated bounds for the survey period 1974-1981 include zero. By the second survey period (1982-1996), the estimated bounds exclude zero, along with the corresponding Imbens and Manski (2004) 95% confidence intervals. The estimated lower bound for this military service effect is 1.1 pp, which implies that the effect is at least 9.2% of the nonveteran mean. For the third and fourth survey periods (1997-2005 and 2006-2013), the estimated lower bounds are considerably higher, consistent with larger detrimental health effects, while the corresponding confidence intervals still exclude zero and their lower endpoint is substantially higher as well. The estimated lower bounds for these two latter survey periods indicate that military service increases the incidence of activity limitation by at least 7.1 pp and 5.8 pp, respectively. These lower bounds imply that the effect is at least 64.2% and 50.3% of the corresponding nonveteran mean. The results in Panel B for the same outcome for nonwhites show a very similar pattern, with the exception that the confidence intervals for the estimated bounds in the survey period 1982-1996 marginally include zero. The magnitudes of the estimated lower bounds for nonwhites in the last two survey periods are 4.3 pp (40.2%), and 11.7 pp (110.9%), respectively.

Consider now Panels E and F in Figure 3, which show the estimated bounds on the incidence of self-reported fair/poor health for whites and nonwhites, respectively. All the estimated bounds for this outcome are positive and exclude zero. However, their 95% confidence intervals do not exclude zero for the first survey period (1974-1981) for both whites and nonwhites, and also for the second and third survey periods (1982-1996 and 1997-2005) for nonwhites. For whites (Panel E), the estimated bounds and confidence intervals exclude zero in the last three survey periods. These estimated lower bounds indicate that military service increases the incidence of fair/poor health by at least 1.1 pp (12.5%), 4.2 pp (48.0%) and 2.7 pp (29.7%), respectively for the 1982-1996, 1997-2005, and 2006-2013 survey periods. For

[36]Recall from section 4.3.3 that the estimated upper bounds correspond to the worst-case bounds. Since these bounds are relatively less informative, we focus our discussion on the estimated lower bounds. Some of our estimated upper bounds, however, can rule out plausible large effects (e.g., those in panel K for current smoker).

nonwhites, in the last survey period, the estimated lower bound indicates that military service increases the incidence of fair/poor health by at least 6.7 pp, or 76.4% and the confidence interval excludes zero.

The outcomes work limitation (panels G and H) and work unable (panels I and J) are likely related to labor force participation and to benefit claims such as VDC. Both of these outcomes—which are only available in the last three survey periods—show similar patterns, which agree with the general pattern of adverse military service health effects described above. Consider the work unable outcome. All the estimated bounds and their 95% confidence intervals exclude zero. The two exceptions are the confidence intervals for nonwhites in the 1982-1996 survey period and, after the multiple testing adjustment, the confidence intervals for whites in the 1982-1996 survey period. For white volunteers (Panel I), the estimated lower bounds for the 1997-2005 and 2006-2013 survey periods indicate that serving in the military increases the incidence of work unable by at least 3.9 pp (63%) and 2.7 pp (38.8%), respectively. For nonwhite volunteers (Panel J), the estimated lower bounds for the last two survey periods indicate that serving in the military increases the incidence of work unable by at least 3.6 pp (64.9%) and 5.9 pp (94.4%), respectively.

The general pattern in the statistical inference on the effects of military service on the general health outcomes of volunteers just described stands in sharp contrast to the point estimates of the same effects for compliers in section 5.3.1. In particular, while the military service effects for compliers do not show a clear pattern and are mostly statistically insignificant, the bounds on the same effect for volunteers are consistent with detrimental and economically significant health impacts from military service, which show up more clearly (in terms of statistical significance) in the long run. Moreover, as will be shown below, the same general pattern for volunteers arises for the military service effects on the group outcomes of activity-limiting chronic conditions and other chronic conditions, reinforcing the notion that these effects are likely not spurious. These new results for volunteers, in turn, also stand in contrast with the estimated military service effects for compliers in the existing literature (e.g., Dobkin and Shabani, 2009; Eisenberg and Rowe, 2009; Angrist et al., 2010).

We turn to the results for the military service effects on the health behaviors of volunteers, which are also shown in Panels K to N of Figure 3. For these outcomes, most of the estimated bounds and their 95% confidence intervals exclude zero, with the exception of the current drinker health behavior for whites in the 2006-2013 survey period (Panel M) and nonwhites in the 1997-2005 survey period (Panel N). Looking at the outcome current smoker for whites (Panel K), the estimated lower bounds indicate that serving in the military increases the incidence of being a current smoker by at least 15.6 pp (50.8%), 12.8 pp (41.9%), 11.9 pp (65%), and 9.5 pp (62.8%) for the survey periods of 1974-1981, 1982-1996, 1997-2005, and 2006-2013, respectively. For nonwhite volunteers (Panel L), the corresponding estimated lower bounds are 13.8 pp (44.9%), 16.8 pp (55.1%), 11.2 pp (62.6%), and 20.9 pp (146%). The results on the military service effects on the incidence of being a current smoker for volunteers stand in stark contrast with the estimates of the same effects for the subgroup of compliers, which are statistically insignificant (Figure 2, Panels K and L). The smoking results for volunteers presented here stand also in contrast to the absence of long-run effects from Vietnam-era military service of compliers documented by Eisenberg and Rowe (2009) for the time period 1997-2005. Conversely, the strong detrimental effect on smoking that we document for volunteer veterans from the Vietnam War is consistent with similar effects documented for veterans from World War II and the Korean War in Bedard and Deschênes (2006).

In the remaining of this section, we briefly discuss the results for the group of outcomes on various activity-limiting chronic conditions and other chronic conditions available in the NHIS. We note that, for these two groups of outcomes, we have had to *conservatively* approximate their corresponding confidence intervals because access to our Research Data Center (RDC) has not been possible due to the Covid-19 pandemic. The appropriate confidence intervals will be replaced once normal access to our RDC is restored.

The estimated bounds on the military service effects on the available activity-limiting chronic conditions are presented in Figure 4 for white volunteers and in Figure 5 for nonwhite volunteers. The general pattern of a detrimental military service effect on health also appears in the results for this set of outcomes. First, none of the 95% confidence intervals of the estimated bounds for either whites or nonwhites in Figure 4 exclude zero for the first survey period of 1974-1981 (up to 8 years after the end of the conflict), except the injuries outcome for white volunteers (Figure 4, Panel 12). For the second survey period (1982-1996; up to 23 years after the end of the conflict), there is only one type of activity-limiting chronic conditions that exhibits a statistically significant effect from military service for white volunteers: musculoskeletal (Panel 13 Figure 4). This type of chronic conditions causing activity limitations is relevant, as the strain of war can lead to musculoskeletal conditions that relate to physical disability (Belmont et al., 2016).[37] The estimated lower bounds indicate that, by the 1982-1996 survey period, military service increases the incidence of musculoskeletal chronic conditions of white volunteers by at least 1.9 pp (29.2%), and of nonwhite volunteers by at least 2.3 pp (37.2 %).

By the 1997-2005 survey period (up to 32 years after the end of the conflict), there are several more activity-limiting chronic conditions that experience statistically significant impacts from military service. The activity-limiting chronic conditions for volunteers in 1997-2005 for which our estimated bounds on the military service effect and their confidence intervals exclude zero (after accounting for multiple testing) are: arthritis, back/neck, and fracture for whites (Figure 4); and diabetes, heart, back/neck, and depression for nonwhites (Figure 5). By the 2006-2013 survey period (up to 40 years after the end of the conflict), the number

[37] For instance, Angrist et al. (2010) document that generalized musculoskeletal conditions is the top 8th condition among Vietnam veterans who were receiving VDC in 2005.

of activity-limiting chronic conditions that are statistically significantly affected by Vietnam-era military service of volunteers largely persist (after conservatively accounting for multiple testing). These conditions are: heart, cancer, back/neck, depression, and fracture for whites (Figure 4); and depression for nonwhites (Figure 5). We single out depression—which is highly intercorrelated and comorbid with PTSD (e.g., Elhai et al., 2011)—to illustrate the magnitude of the effects in the last two survey periods. For white volunteers, Panel 18 in Figure 4 indicates that military service increases the incidence of activity-limiting depression by at least 2.3 pp (345%) for the 1997-2005 survey period (which is not statistically significant), and by at least 2.7 pp (375%) for the 2006-2013 survey period. Meanwhile, for nonwhite volunteers, Panel 18 in Figure 5 indicates that the corresponding estimated lower bounds are 2.3 pp (1097%) and 4.5 pp (1146%).[38]

Last, we briefly discuss the results for the group of outcomes related to other chronic conditions, which are only available in the NHIS 1997-2005 and 2006-2013 survey periods. Due to this limited availability, we are not able to compare the evolution of the military service effect since the end of the conflict. Nevertheless, given the general pattern that has been uncovered throughout this section, we expected to find statistically significant effects on various of these other chronic conditions for volunteers. This is exactly what we report. The estimated bounds for the military service effect on other chronic conditions of volunteers are presented in Figure 6 for whites and Figure 7 for nonwhites. Nine of the twenty-five other chronic conditions available in the 1997-2005 survey period are statistically significantly affected by Vietnam-era military service for white volunteers (Figure 6). The conditions with statistically significant effects that withstand our conservative multiple testing adjustment are emphysema, feeling interferes with life, hearing conditions, severe hearing conditions, joints, liver, neck pain, lower back pain, and (loss of all lower and upper) teeth. For the 2006-2013 survey period, the following chronic conditions are statistically significantly affected by military service (and withstand our multiple testing adjustment) for white volunteers: cancer, hearing conditions, severe hearing conditions, and teeth.

For nonwhite volunteers (Figure 7), the evidence is similar. In the 1997-2005 survey period, four of the twenty-five other chronic conditions are statistically significant and withstand our multiple testing adjustment: hypertension, neck pain, lower back pain, and ulcer. For the 2006-2013 survey period, more chronic conditions are statistically significantly affected by military service (and withstand our multiple testing adjustment): cancer, diabetes, emphysema, headache, heart attack, joints, neck pain, lower back pain, sinus, teeth, ulcer, and worse health compared with 12 months ago. To provide an illustration of the magnitude of the effects, con-

[38]The percentage increases relative to the non-veteran mean implied by the estimates on activity-limiting depression may seem large. We note that they are of a similar order of magnitude as the raw differences in incidence of activity-limiting depression between veterans and nonveterans shown in Tables A3 and A5 in the Online Appendix.

sider the chronic conditions related to lower back pain (in Panel 19 of Figures 6 and 7), which may lead to mobility limitations and potentially to claims for disability compensation. Our estimated bounds and corresponding confidence intervals that exclude a zero military service effect on the incidence of lower back conditions for volunteers indicate an impact of at least 7.1 pp (26.2%) and 5 pp (18.4%) for whites in the 1997-2005 and 2006-2013 survey periods (the latter being only marginally significant after our multiple testing adjustment). Similarly, for nonwhite volunteers, the estimated lower bounds indicate an impact of at least 9.3 pp (36.6 %) and 19.1 pp (76.3 %), in the corresponding last two survey periods.

To summarize the results for the subgroup of volunteers, our estimated bounds on the Vietnam-era military service effect for this group uncovers statistically and economically significant detrimental effects on a myriad of health outcomes, particularly in the survey periods of 1997-2005 and 2006-2013, that is, about 24 to 40 years after the end of the Vietnam War.

5.3.4 The Effects of Military Service on Health Outcomes and Behaviors of the Population of Veterans

The entire group of Vietnam-era veterans is a policy-relevant population. Policies related to veterans do not distinguish by latent subpopulations of volunteers or compliers. Still, to our knowledge, there are no studies providing causal statistical inference for the population of veterans while exploiting the Vietnam-era draft lotteries as source of exogenous variation. We present results for this population here. Since the population of veterans consists of the latent subpopulations of compliers and volunteers, we essentially combine the point estimates for compliers in section 5.3.1 with the estimated bounds for volunteers in section 5.3.3, weighted by their corresponding proportions in the population of veterans. Given that the estimated proportions indicate that volunteers represent about three-quarters of all veterans, it is not surprising that the results for veterans show the same pattern as the results for volunteers. That is, the results for veterans imply statistically and economically significant detrimental effects of military service on their health outcomes and behaviors, particularly in the long run. We briefly discuss the main findings for veterans in this section, while the complete set of results can be found in Online Appendix A, Tables A.16 to A.18.

Tables A.16-A.17 in Online Appendix A present the estimated bounds for white and nonwhite veterans on their general health outcomes and health behaviors. The results for the group of general health outcomes show the same general pattern as for volunteers: the first survey period (1974-1981) does not exhibit statistical significant military service effects, while the effects for most general health outcomes in subsequent survey periods are statistically significant (for both white and nonwhite veterans). For illustration, consider activity limitation. The estimated lower bounds on the military service effect with confidence intervals that exclude zero—and withstand our conservative multiple testing adjustment—indicate increases in the incidence of activity limitation of white veterans by at least 1.7 pp (14.7%), 4 pp (36.4%), and 4.2 pp (36.7%) for the survey periods 1982-1996, 1997-2005, and 2006-2013, respectively. The corresponding estimated lower bounds for nonwhite veterans for the same general health outcome and survey periods are 3.5 pp (32%), 4 pp (37.1%), and 9.1 pp (86.5%).

The results for the military service effect on the health behaviors of the population of veterans (Online Appendix A, Tables A.16-A.17) indicate statistically significant detrimental impacts in all survey periods (with the exception of current drinker in NHIS 1997-2005 and 2006-2013 for whites and the last survey period for nonwhites). This pattern is similar to the pattern found for volunteers, with the estimated lower bounds for veterans usually being somewhat smaller. Considering white veterans for an illustration, their estimated lower bounds in Table A.16 and Table A.17 indicate military service effects that increase the incidence of being a current smoker by at least 12.2 pp (40%), 9.8 pp (31.8%), 8 pp (44.1%), and 6.4 pp (42.6%) for the respective survey periods in chronological order.

The results for the military service effect on the activity-limiting chronic conditions and other chronic conditions of veterans are presented in Online Appendix A, Tables A.16 to A.18. By in large, the military service effects on chronic conditions that were statistically significant for volunteers are also statistically significant for veterans. However, in the case of veterans, a fair number of these effects are no longer statistically significant after applying our conservative multiple testing procedure.[39] For white veterans, the military service effects on chronic conditions that remain statistically significant are: musculoskeletal in the 1982-1996 survey period, back/neck in the 1997-2005 survey period, and back/neck, depression, and hearing conditions in the 2006-2013 survey period. In contrast, for nonwhite veterans, none of the military service effects on chronic conditions remain statistically significant after the multiple testing adjustment.

In summary, the results about the military service effects on health outcomes and behaviors for the population of Vietnam-era veterans indicates a general pattern of statistically insignificant effects in the short run, followed by statistically significant impacts of non-negligible magnitude in the long run (24 to 40 years after the end of the conflict). This pattern mirrors the one for the latent subgroup of volunteers, who represent the bulk of the Vietnam-era veterans.

5.4 Mortality Outcomes

Using the NHIS mortality linkage 1985-2009, we analyze the effects of the draft lotteries and military service on mortality in three survey periods that correspond to different survey structures: 1985-1996, 1997-2005, and 2006-2009. Analyzing mortality serves two purposes.

[39] Although conservative, we prefer to focus on the effects that attain statistical significance after the multiple testing adjustment. Nevertheless, as mentioned before, the confidence intervals were conservatively approximated for these chronic conditions. Thus, once they are computed at the RDC, it is possible that more effects on chronic conditions for veterans achieve statistical significance after the multiple testing adjustment.

First, it allows us to assess an additional effect that military service can have on an important health outcome in mortality. Second, the analysis of the effects of the draft lotteries and military service on mortality provides an opportunity to assess the extent by which our earlier results could be impacted by differential mortality among the groups under study (draft eligible/ineligible, veterans/nonveterans, compliers/volunteers).

The mortality outcome considered here is an indicator for whether the individual has died by December 31, 2011, when males in the affected cohorts would be between 59 and 63 years old. Given this outcome, our discussion below focuses mostly on the effects for the first survey period (1985-1996) for two reasons. First, this is the period with the largest gap between the time the individuals are surveyed and the time mortality is last measured in our data. Second, it is the earliest period available in our data, and thus the one whose sample composition is least likely to have been affected by mortality related to the draft lotteries or military service.[40] Still, for completeness, we also present the results on mortality for the two subsequent survey periods available.

Table 5 presents the results for the mortality outcome. The two vertical panels in the table refer to whites and nonwhites, while the three columns in each vertical panel indicate the survey periods under consideration. The first and fourth columns refer to the first survey period (1985-1996). The first row provides the effect of the draft eligibility status on mortality by the end of 2011 (the ITT). The estimates indicate that draft eligibility status does not impact mortality, with all the point estimates small and statistically insignificant. These results are different from those in Hearst et al. (1986), who find that draft eligible men had a statistically significant 4% higher mortality rate than non-eligible men using death certificate records from California and Pennsylvania between 1974 and 1983. However, our results are consistent with more recent work that fails to find evidence of an effect of the Vietnam-era draft lotteries on mortality, such as Angrist et al. (1996), who use a similar sample to that of Hearst et al. (1986); Conley and Heerwig (2009), who use a larger national sample of death records from 1989-2002; Eisenberg and Rowe (2009), who use 1978-1980 and 1997-2005 NHIS samples; and Angrist et al. (2010), who use the 2000 Census. Our results are also consistent with the lack of evidence of an effect of draft eligibility on mortality in the NHIS 1974-2013 data used in the previous sections. If draft eligibility affected mortality, we would expect the fraction of draft-eligible men to change over time. However, as shown in Table 1, the fraction of draft-eligible men by veteran status and race is stable over the four survey periods (the same occurs when dividing the sample only by race, which is not shown in the table), consistent with the absence of an effect of draft eligibility on mortality.

The second row shows OLS estimates of the effect of military service on mortality by the end of 2011. The OLS estimate for whites in the 1985-1996 survey period indicates that

[40]Unfortunately, the mortality linkage is not available prior to 1985.

serving in the military is associated with a statistically significant 1.2 pp increase in mortality by 2011 (corresponding to a 13.3% increase relative to the nonveterans' mortality rate of 9 pp), while the corresponding estimate for nonwhites is close to zero and not statistically significant. 2SLS estimates that use draft eligibility as an IV to estimate the average treatment effect of military service on mortality for compliers are presented in the third row. None of these estimated impacts is statistically significant, suggesting that compliers do not exhibit an impact of military service on their mortality by 2011. These results are in line with those in Angrist et al. (1996). Hearst et al. (1986), however, found small statistically significant effects on mortality for compliers that were mainly driven by suicide and motor vehicle accidents. Bedard and Deschênes (2006) also found significant effects of military service on mortality for World War II and Korean War veterans using instrumental variables other than the Vietnam-era draft lotteries.

The previous-to-last row presents estimated bounds on the effect of military service on the mortality of volunteers. Focusing on the 1985-1996 survey period, for both whites and nonwhites, the estimated bounds are positive and exclude zero. For whites, the 95% confidence interval also excludes zero. According to the estimated lower bound, this statistically positive impact indicates that, for white males in the 1985-1996 survey period, military service increases the probability of dying by the end of 2011 by at least 2.2 pp, a 24.4% increase relative to the nonveterans' mortality rate. The last row indicates that the estimated bounds on the effect of military service on the mortality of the population of veterans behave similar to those of volunteers. For white veterans, the estimated bounds in the 1985-1996 survey period again exhibit a statistically positive (at the 95% level) effect of military service on mortality, which is at least 1.2 pp (13.3%).[41]

In summary, based on the 1985-1996 survey period, the results on the outcome of mortality by 2011 in Table 5 suggest that military service has a statistically positive effect on the mortality of volunteers and veterans, while no such effect is detected for compliers. These findings are consistent with our previous results of deteriorating health effects of military service for volunteers and veterans, and no such effects for compliers. In passing, the results in Table 5 also illustrate that finding no effect of the draft lotteries on mortality does not imply that there are no effects of military service on mortality, which is sometimes overlooked in the literature.

In light of the results pertaining to mortality in this section, a discussion on their possible implications for our results in section 5.3 is appropriate. The presence of a military service effect on mortality does not lead to "survivor bias" in our estimates of the effects of military

[41] The results for the effect of military service on the probability of dying by 2011 for volunteers and veterans in the two most recent survey periods that are closer to 2011 suggest that these effects may be smaller (as one would expect), since the estimated bounds are smaller in magnitude and do not statistically rule out a zero effect.

service on the health outcomes and behaviors in section 5.3, as long as draft eligibility does not affect mortality (which is consistent with the evidence above). The reason is that, under assumptions A1 to A6, the latent subpopulations of compliers, volunteers, and never takers in our data are still comparable——or unaffected by differential mortality due to military service––under $Z = 0$ and $Z = 1$, allowing us to estimate the quantities needed to make inference on the effects of interest (i.e., $LATE_c$, $LATE_{at}$, and ATT) without survivor bias. Specifically, for compliers (for whom $Z = D$), having no effect of draft eligibility on mortality implies that the groups of compliers under $Z = 0$ and $Z = 1$ remain statistically comparable. For volunteers (respectively, never takers), all (none) of them serve in the military, so under the exclusion restriction their mortality is statistically the same under $Z = 0$ and $Z = 1$.[42] Nevertheless, the presence of a military service effect on mortality does change the interpretation of the parameter for which the effects are estimated over the different survey periods, as the different NHIS survey periods may no longer be comparable. In particular, under the presence of a military service effect on mortality, the effects on health outcomes and behaviors in the different survey periods are estimated without survivor bias for those individuals who are alive at least up to the survey period at which the estimation takes place. This is a policy-relevant population, as this is the group of individuals experiencing the health effects of military service during that survey period (since the other men have already died). Furthermore, regarding the health results on volunteers and veterans in sections 5.3.3 and 5.3.4, note that if military service increases mortality for these groups, and assuming that the less healthy die over time at a higher rate than the healthy (as one would expect), our estimated lower bounds on the long-run health effects of military service for later survey periods can be seen as conservative for those men who survived up until the earlier survey periods. For example, the estimated lower bounds on the health effects of military service for veterans in the 2006-2013 survey period (which are unbiased for the effects for veterans who had lived at least up to 2006-2013 if draft eligibility does not affect mortality) can be seen as conservative for the effects for veterans who had lived at least up to the 1974-1981 survey period.[43,44]

[42]There would be survivor bias when estimating the health effects for volunteers given our finding of an effect of military service on mortality for them if, for instance, we were comparing the health of volunteers who served ($D = 1$) to the health of volunteers who did not serve ($D = 0$), but that is not how our bounds are constructed, as there are no volunteers with $D = 0$ in our data.

[43]Even if, contrary to the evidence above, there was a positive effect of draft eligibility on mortality, and assuming the less healthy die over time at a higher rate than the healthy (as one would expect), this would actually make it harder to detect long-term detrimental health effects from military service. In this case, our estimated long-term health effects of military service for compliers and our estimated lower bounds for the population of veterans would be seen as conservative. As for the estimated bounds on the effects for volunteers, these bounds would remain unbiased for the effect for volunteers who had lived up to the survey period at which estimation takes place, since the quantities needed to construct these bounds are unaffected by differential mortality due to draft eligibility under the exclusion restriction (see equations (5) and (6)).

[44]It is possible to derive bounds on the average effect of military service on the health outcomes and behaviors for the same specific latent subpopulation over time (e.g., for compliers, volunteers, or veterans who would

Lastly, there is a potential threat to assumption A6 if (for any reason) there exists differential mortality between never takers and volunteers, had both groups not served. Concretely, A6 would be violated if unhealthy never takers were to die at a faster rate than unhealthy volunteers—had both subpopulations not served—in such a way that the direction of the inequality in A6 flipped. Under this differential mortality, we would expect the average health of never takers to improve over time (in absolute or relative terms). However, this is not the case in our data, even relative to the average health of compliers in the absence of service (i.e., $E[Y(0)|c]$), which is point identified (results not shown in tables). In addition, from the analysis of pre-draft characteristics in Table 4, the pre-draft difference in average health (as measured by activity limitations before 1965) between never takers and volunteers seems too large to be totally eliminated by differential mortality between these two groups (in the absence of military service). Therefore, we find it highly unlikely that our results could be driven by this potential threat to assumption A6.

6 Discussion

6.1 Some Implications of the Results

Previous studies that estimated the military service health effect on veteran compliers using the draft lotteries as an IV (e.g., Dobkin and Shabani, 2009, Eisenberg and Rowe, 2009, Angrist et al. 2010, Schmitz and Conley, 2016) largely find no statistically significant effect in the short or long run. The results presented in the previous section for compliers using the same identification strategy agree with those studies. In contrast, our results indicate that Vietnam-era volunteers and the entire population of veterans experience detrimental effects of military service on a myriad of health outcomes and behaviors, which starts to unveil after 10 years of the end of the conflict and appear strongest in the long run (24 to 40 years after the end of the Vietnam War). These findings are relevant to studies that have attributed the recent rise in VDC compensation costs to the policy change in 2001 that allowed "boots on the ground" Vietnam veterans with diabetes mellitus to qualify for VDC benefits (Singleton, 2009, Autor et al., 2011), and the increased post-traumatic stress disorder (PTSD) diagnosis among the Vietnam veterans (Autor et al., 2011). By showing that Vietnam-era military service statistically significantly causes higher activity-limiting depression (which is highly intercorrelated and comorbid with PTSD) and diabetes for both white and nonwhite volunteers and veterans, our results suggest that the increased VDC compensation might have indeed been induced by Vietnam-era veterans who suffer from these chronic conditions.

be alive by the end of 2011 regardless of their eligibility to draft) taking into account that the individuals in the different NHIS survey periods may not be comparable. This can be done in the spirit of the work by Lee (2009) and Chen and Flores (2015). These bounds would remain valid even in the presence of an effect of draft eligibility on mortality. We consider such analysis, however, beyond the scope of this paper.

We provide an idea of the implied incremental costs of the detrimental effects on depression and diabetes for Vietnam-era veterans in the 1948-1952 cohorts. To do this, we use the per-person estimated direct costs of depression for 2010 in Greenberg et al. (2015) of \$5,061 and the estimated direct costs of diabetes for 2012 in American Diabetes Association (2013) of \$7,900. We estimate the number of (white and nonwhite) Vietnam-era veterans in the 1948-1952 cohorts by using the Vital Statistics of the U.S. and our estimated proportion of veterans from the NHIS 2006-2013 survey period. This yields a total of 2,390,146 white and 305,681 nonwhite veterans from those cohorts. Multiplying the number of veterans by the corresponding estimated *lower bounds* for depression and diabetes yields increased incidences of *at least* 62,861 (whites) and 12,930 (nonwhites) for depression, and 118,790 (whites) and 1,620 (nonwhites) for diabetes. Using the cost per person, the increased annual costs are, *at least*, \$384 million for depression and \$951 million for diabetes. The Department of Veteran Affairs (2020) reports \$42.4 billion in total 2010 expenditures on medical care by the VA. Thus, for 2010, the approximate increased medical costs for Vietnam-era veterans in the 1948-1952 cohorts from only these two conditions correspond to *at least* 3.15% of that year's total medical expenditures by the VA. For reference, we calculate that Vietnam-era veterans in the 1948-1952 cohorts make up about 12% of all veterans in 2010. In sum, the detrimental health effects we uncover appear economically non-negligible.[45]

Angrist et al. (2010) documented the most common diagnoses for Vietnam veterans collecting VDC claims. A number of those diagnoses coincide with reported health conditions for which we have found statistically significant detrimental effects of military service for either volunteers or the population of veterans. Those diagnoses are: diabetes, PTSD (of which depression—our measure—is highly comorbid), hearing conditions, and generalized musculoskeletal conditions. Considering musculoskeletal conditions (e.g., lower back and neck conditions) as one of the conditions that strongly and consistently appear as adversely affected by military service in our results, Angrist et al. (2010) documented that their diagnosis is in the top-9 of conditions among Vietnam veterans who were receiving VDC in 2005. This is consistent with studies that document a detrimental association between military service and musculoskeleteal conditions on U.S. veterans (Hinojosa and Hinojosa, 2016) and on Canadian veterans (Thompson et al., 2015). Moreover, there is evidence that U.S. veterans with musculoskeletal conditions tend to exhibit higher rates of opioid dosing (Han et al, 2017), and anxiety and substance use disorders (Patten et al, 2006). Therefore, policies and resources towards treating and controlling musculoskeletal conditions may indirectly help to reduce other mental conditions and substance abuse among veterans.

[45] One can also compare the increased medical costs from these two conditions to the corresponding total direct costs for each type of conditions in the US. Using the same corresponding sources as for the per-person estimates, the total direct costs are \$77.9 billion for depression and \$176 billion for diabetes. Thus, the increased medical costs for Vietnam-era veterans in the 1948-1952 cohorts correspond to *at least* 0.49% and 0.54% of the total annual US direct costs for depression and diabetes, respectively.

The results on the military service effect on volunteers and veterans also suggest strong health detrimental impacts on smoking and (to a lesser extent) drinking. These findings for smoking are consistent with the findings in Bedard and Deschênes (2006) who studied World War II veterans, and with those in London et al. (2017) who examined the effect of general active duty during military service across several cohorts of U.S. men. The negative military service effect on health behaviors may represent a plausible mediator for the detrimental long-term effects of military service on certain chronic conditions of veterans, such as emphysema, hypertension, cancer, and heart conditions. This finding is particularly relevant to recent policies increasing the price of cigarette products in U.S. military bases following legislation from 2016 (Jowers, 2016) and the planned increase of the minimum smoking age to 21 among U.S. troops (Shane, 2019). These policies appear promising to curb cigarettes smoking in the military and possibly lead to long-term improvements on the veterans' health and cost savings. In addition, our results suggest that resources allocated to special smoking cessation programs for veterans (e.g., Lowry, 2016) could be beneficial, as they can potentially curb the detrimental long-term health effects of veterans from military service.

6.2 Why Could the Results for Compliers and Volunteers Be Different?

Our results suggest that the effect of Vietnam-era military service on health outcomes and behaviors differs between compliers and volunteers: the former subgroup does not seem to display detrimental health effects, while volunteers exhibit strong detrimental health effects in the long run. One of our previous exercises suggests a factor that may contribute to the contrasting results between these two latent subgroups. The average pre-draft characteristics presented in Table 4 suggest that white compliers had a statistically significant lower high school incompletion rate by about 4.5 pp relative to their volunteer counterparts. For nonwhites the same pattern holds, although the estimated differences are noisy. In light of this evidence, higher educational attainment might have helped compliers obtain additional college-level education and degrees after military service through the GI Bill. This would be consistent with the findings in Angrist and Chen (2011) that Vietnam-era military service particularly increased college attendance and associate and bachelor's degree attainment among compliers.[46] Given the positive correlation between educational attainment and health (e.g., Zajacova, 2006; Oreopoulos and Salvanes, 2011), and the finding that the detrimental health effects of military service appear in the long run, it is possible that higher educational attainment after service may have played a role in partly offsetting some of the detrimental health effects of military

[46] Angrist and Chen (2011) also found comparatively smaller effects on the high school completion rate of compliers (about 2 pp), which they conjecture are due to complier veterans obtaining GEDs or receiving honorary high school diplomas.

service for compliers. Moreover, the higher educational attainment of compliers might have helped them better adapt to civilian lives after military service over time, further contributing to the alleviation of the detrimental impact of military service on long-term health.[47] Other factors may have also contributed to the different impacts of Vietnam-era military service on the long-term health of compliers and volunteers. For instance, it may had been the case that volunteers served longer and were more likely to be exposed to combat than compliers, resulting in stronger detrimental health effects in the long run. Future research would benefit from analyzing in detail the possible sources of such heterogeneous effects.

7 Conclusion

This paper analyzed the effect of the U.S. Vietnam-era military service on veterans' health outcomes and behaviors using a restricted version of the National Health Interview Surveys (NHIS) 1974-2013. These data contain a comprehensive list of health outcomes and behaviors, and allow an analysis of the evolution of the health effects of military service over different periods of time after the end of the Vietnam conflict. To overcome the selection-into-military-service problem, we employ the Vietnam-era draft lotteries as an instrumental variable (IV) for military service, following prior literature that estimates effects for the latent subgroup of compliers (Angrist et al., 1996; Dobkin and Shabani, 2009; Angrist et al., 2010), who comprise less than 25% of all Vietnam-era veterans. In contrast to this prior literature, we further leverage the draft lotteries IV and a relatively mild mean weak monotonicity assumption to estimate sharp nonparametric bounds on the military service effect for volunteers (those who enlist regardless of their eligibility to draft) and for the entire population of veterans. The mean weak monotonicity assumption employed states that, in the absence of military service for both latent groups, the average potential health outcomes of the never takers (those who do not serve regardless of their eligibility to draft) are no better than those of the volunteers. We argue that this assumption is plausible in light of the physical, medical, and mental tests performed by the U.S. military before enlisting individuals, which results in an almost-mechanic positive health selection into the military. We also present indirect statistical evidence supporting the plausibility of this assumption. More generally, this exercise illustrates how a generally accepted notion—in this case the positive self-selection into the military based on health (e.g., Seltzer and Jablon, 1974; Bedard and Deschênes, 2006; Eisenberg and Rowe, 2009)—can be statistically formalized to construct and estimate informative bounds, which can be useful in other empirical settings.

The key findings in this paper are three. First, we assess the validity of the draft lotteries

[47]Note that, as also pointed out in Angrist and Chen (2011), effects of military service on educational attainment (e.g., via the GI Bill) do not imply a violation of the exclusion restriction. The same reasoning applies to Assumption A6.

IV by quantifying, for the latent subpopulation of never takers, the net effect of the draft lotteries on the health outcomes and behaviors that operates through channels other than military service (e.g., through draft avoidance behaviors). In line with prior studies focusing on labor outcomes (Kitagawa, 2005; Mourifié and Wan, 2017), we find no statistical evidence of invalidity of the IV for any of the health outcomes and behaviors we consider. Second, our point estimates of the Vietnam-era military service effect on the health outcomes and behaviors of complier veterans do not provide consistent evidence of statistically significant effects. This is also in line with prior literature using the draft lotteries as an IV for military service (e.g., Dobkin and Shabani, 2009; Angrist et al., 2010).

Third, for Vietnam-era volunteers, who account for over 75% of all veterans, their estimated bounds show clear statistically significant detrimental health effects of important magnitude that seem to appear in the long term. In the approximately eight years after the end of the conflict (the 1974-1981 survey period), no military service effects are statistically significant (i.e., no estimated bounds and their confidence intervals exclude zero) for any health outcomes, with the exception of the health behavior of smoking. The estimated bounds, separately estimated for whites and nonwhites, are consistent with an effect on smoking of at least 14 pp and at most 46 pp (44.9% to 151% of the nonveteran mean). By the second survey period—nine to twenty-three years after the end of the conflict—the health of white volunteers shows statistically significant detrimental impacts of military service on activity and work limitation, self-reported fair/poor health, muscoskeletal conditions, and smoking. For nonwhite volunteers, in the second survey period we find statistically significant detrimental military service effects on musculoskeletal conditions and smoking. For the last two survey periods of 1997-2005 and 2006-2013—twenty-four to forty years after the end of the conflict—we find statistically significant detrimental military service effects for both white and nonwhite volunteers on a considerable number of health outcomes and behaviors. To name a few, military service increases the incidence of both white and nonwhite volunteers on activity and work limitation, self-reported fair/poor health, musculoskeletal conditions, and the smoking. Military service also increases their incidence on a myriad of chronic conditions, for example, back/neck conditions, lower back conditions, depression and diabetes.[48] To give a sense of the magnitude of the estimated effects, consider the health outcome activity limitation for the last two survey periods. The estimated lower bounds for white volunteers indicate that military service increases activity limitation by at least 7.1 pp (64.2% of the non-veteran mean) and 5.8 pp (50.3%) in the last two survey periods, respectively. For nonwhites, their estimated lower bounds indicate that the same effect is at least 4.3 pp (40.2%) and 11.7 pp (110.9%) for the last two survey periods, respectively. The previous set of results, perhaps not surprisingly, imply that the military service effects on the entire population of Vietnam-era veterans largely

[48]Throughout our analysis, we employ a conservative multiple testing procedure to adjust the statistical inference of groups of multiple health outcomes and behaviors.

follows the trends of the volunteers (albeit with a smaller number of statistically significant effects), as volunteers account for most of the veterans. We also document results on the Vietnam-era military service impact on the mortality of veterans by 2011, and discuss their implications for the previous findings about the military service impact on health outcomes and behaviors.

Our findings imply that there have been detrimental health effects of Vietnam-era military service, which show up and appear to worsen over time. They also suggest that these detrimental effects are primarily driven by males who volunteered for enlistment, and not by compliers. These results can have relevant implications for policies regarding compensation of veterans after service, for which prior studies have documented increases in transfers due to various reasons that relate to health outcomes (e.g., Angrist et al., 2010; Singleton, 2009; Autor et al., 2011). By documenting that the Vietnam-era military service statistically—and economically—significantly worsened long-run general health outcomes, depression, and musculoskeletal conditions, among other chronic conditions, our results increase our understanding of the steep increase in payments from the VDC to Vietnam veterans. Lastly, our results documenting the smoking-inducing effect of military service among Vietnam-era veterans (driven by volunteers), suggest that this behavior could be a channel for the delayed appearance of some of the detrimental health effects of military service. And that recent policies to curve smoking within the military (e.g., increases in cigarette prices on military bases and of the minimum smoking age) may help to improve veterans' healthy behaviors, thereby reducing the healthcare cost for veterans in the long-run.

References

[1] Almond, D. and Currie, J. (2010) *Human Capital Development before Age Five*. 1st ed. Cambridge, MA.: National Bureau of Economic Research.

[2] American Diabetes Association (2013) Economic Costs of Diabetes in the U.S. in 2012. *Diabetes Care*, 36(4): pp.1033-1046.

[3] Angrist J. (1990) Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records. *American Economic Review*, 80 (3): pp.313-336.

[4] Angrist, J. (1991). The Draft Lottery and Voluntary Enlistment in the Vietnam Era. *Journal of the American Statistical Association*, 86(415), p.584-595.

[5] Angrist, J. (2004). Treatment Effect Heterogeneity in Theory and Practice. *Economic Journal*, 114: pp.C52–83.

[6] Angrist, D., and Chen, S. (2011). Schooling And The Vietnam-Era GI Bill: Evidence From The Draft Lottery. *American Economic Journal: Applied Economics*, 3 (2): pp. 96-118.

[7] Angrist, J., Chen, S., and Frandsen, B. (2010) Did Vietnam Veterans Get Sicker in the 1990s? The Complicated Effects of Military Service on Self-reported Health. *Journal of Public Economics*, 94(11-12): pp.824-837.

[8] Angrist J., Imbens, G., and Rubin, D. (1996) Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91(434): pp.444-472.

[9] Autor, D., Duggan, M. and Lyle, D. (2011) Battle Scars? The Puzzling Decline in Employment and Rise in Disability Receipt among Vietnam Era Veterans. *American Economic Review*, 101(3): pp.339-344.

[10] Bailey, M., and Chyn, E. (2020) The Demographic Effects of Dodging the Vietnam Draft. *American Economic Review*, 110(5): pp.220-225.

[11] Banks, J., Oldfield, Z., and Smith, J. (2012) Childhood Health and Differences in Late-Life Health Outcomes between England and the United States. *Investigations in the Economics of Aging*. Chicago, IL: University of Chicago Press.

[12] Baskir, L., and Strauss, W. (1978) *Chance and Circumstance: The Draft, the War, and the Vietnam Generation*. New York, NY: Alfred A. Knopf.

[13] Bedard, K., and Deschênes, O. (2006) The Long-Term Impact Of Military Service On Health: Evidence From World War II And Korean War Veterans. *American Economic Review*, 96 (1): 176-194.

[14] Belmont, P., Owens, B., and Schoenfeld, A. (2016). Musculoskeletal Injuries in Iraq and Afghanistan: Epidemiology and Outcomes Following a Decade of War *Journal of the American Academy of Orthopaedic Surgeons*, 24 (6): pp. 341-348.

[15] Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate. *Journal of the Royal Statistical Society*, 57: pp. 289-300.

[16] Blanco, G., Chen, X., Flores, C. A., and Flores-Lagunes, A. (2020) Bounds on Average and Quantile Treatment Effects on Duration Outcomes Under Censoring, Selection, and Noncompliance. *Journal of Business and Economic Statistics*, 38(4): 901-920.

[17] Card, D. and Lemieux, T. (2001) Going to College to Avoid the Draft: The Unintended Legacy of the Vietnam War. *American Economic Review*, 91(2): pp. 97-102.

[18] Card, D. and Lemieux, T. (2002) Did Draft Avoidance Raise College Attendance During the Vietnam War? Center for Labor Economics Working Paper No. 46.

[19] Carter, S., Smith, A. and Wojtaszek, C. (2017). Who Will Fight? The All-Volunteer Army after 9/11. *American Economic Review*, 107(5), pp.415-419.

[20] Chen, X. and Flores, C. A. (2015) Bounds on Treatment Effects in the Presence of Sample Selection and Noncompliance: The Wage Effects of Job Corps. *Journal of Business and Economic Statistics*, 33(4): 523–40.

[21] Chen, X., Flores, C. A., and Flores-Lagunes, A. (2018) Going beyond LATE: Bounding Average Treatment Effects of Job Corps Training. *Journal of Human Resources*, 53 (4): 1050-1099.

[22] Chen, X., Flores, C. A., and Flores-Lagunes, A. (2020) Bounds on Average Treatment Effects with an Invalid Instrument: An Application to the Oregon Health Insurance Experiment. Unpublished Manuscript, California Polytechnic State University at San Luis Obispo.

[23] Code of Federal Regulations. (1968) 32 CFR §§1622.2, 1623.2

[24] Congressional Budget Office. (2014). Veterans' Disability Compensation: Trends and Policy Options. [online] Available at: `https://www.cbo.gov/publication/45615` [Accessed 11 Feb. 2018].

[25] Conley, D. and Heerwig, J. (2012) The Long-Term Effects of Military Conscription on Mortality: Estimates from the Vietnam-era Draft Lottery. *Demography*, 49: pp. 841-855.

[26] Davies, P., Purcell, P. and Engelhardt, G. (2015) Vietnam-era Military Service and DI Participation. *Economics Letters*, 136: pp. 5-8.

[27] Department of Veterans Affairs (2020), National Center for Veteran Analysis and Statistics, Expenditures. [online] Available at: `https://www.va.gov/vetdata/expenditures.asp` [Accessed 15 Nov. 2020].

[28] Dobkin, C., and Shabani, R. (2009) The Health Effects of Military Service: Evidence from the Vietnam Draft. *Economic Inquiry*, 47 (1): pp. 69-80.

[29] Eisenberg, D. and Rowe, B. (2009) The Effect of Smoking in Young Adulthood on Smoking Later in Life: Evidence based on the Vietnam Era Draft Lottery. *Forum for Health Economics & Policy*, 12(2).

[30] Elhai, J., Contractor, A., Palmieri, P., Forbes, D. and Richardson, D. (2011) Exploring the Relationship Between Underlying Dimensions of Posttraumatic Stress Disorder and Depression in a National, Trauma-Exposed Military Sample. *Journal of Affective Disorders*, 133: 477-480.

[31] Fienberg, S. (1971) Randomization and Social Affairs: The 1970 Draft Lottery. *Science*, 171 (3968): pp. 255-261.

[32] Fisher, A. (1969) The Cost of the Draft and the Cost of Ending the Draft. *American Economic Review*, 59: pp. 239-254.

[33] Flores, C., and Flores-Lagunes, A. (2010) Nonparametric Partial Identification of Causal Net and Mechanism Average Treatment Effects. Unpublished manuscript, California Polytechnic State University at San Luis Obispo.

[34] Flores, C., and Flores-Lagunes, A. (2013) Partial Identification of Local Average Treatment Effects with an Invalid Instrument. *Journal of Business and Economic Statistics*, 31 (4): pp. 534-545.

[35] Frangakis, C., and Rubin, D. (2002) Principal Stratification in Causal Inference. *Biometrics*, 58, 21-29.

[36] Greenberg, P., Fournier, A., Sisitsky, T., Pike, C. and Kessler, R. (2015) The Economic Burden of Adults with Major Depressive Disorder in the United States (2005 and 2010). *Journal of Clinical Psychiatry*, 76:2 155-162.

[37] Grossman, M. (1972) On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80: pp. 223-255.

[38] Gustman, A., Steinmeier, T. and Tabatabai, N. (2016). Declining Wealth and Work among Male Veterans in the Health and Retirement Study. *The Journal of Retirement*, 4(1), pp.77-98.

[39] Han, L., Allore, H., Goulet, J., Bathulapali, H., Skanderson, M., Brandt, C., Haskell, S., Krebs, E. (2017) Opioid Dosing Trends Over Eight Years Among US Veterans with Musculoskeletal Disorders after Returning from Service in Support of Recent Conflicts. *Annals of Epidemiology*, 27(9), pp. 563-369.

[40] Hanson, D. (2013). *Two Essays on the Interaction between Marriage and Policy.* Boston College, PhD dissertation.

[41] Hearst, N., Newman, T. and Hulley, S. (1986) Delayed Effects of the Military Draft on Mortality. *New England Journal of Medicine*, 314(10): pp. 620-624.

[42] Hinojosa, R. and Hinojosa, M. (2016) Activity-Limiting Musculoskeletal Conditions in US Veterans Compared to Non-Veterans: Results from the 2013 National Health Interview Survey, 11(12): e0167143.

[43] Huber, M., and Mellace, G. (2015), Testing Instrument Validity for $LATE$ Identification Based on Inequality Moment Constraints. *Review of Economics and Statistics*, 97 (2), 398–411.

[44] Imbens, G., and Angrist, J. (1994) Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62 (2): pp. 467-475.

[45] Imbens, G., and Manski, J. (2004) Confidence Intervals for Partially Identified Parameters. *Econometrica*, 72 (6): 1845-1857.

[46] Imbens, G., and Rubin, D. (1997) Estimating Outcome Distributions for Compliers in Instrumental Variable Models. *Review of Economic Studies*, 64 (4): 555–574.

[47] Jowers, K. (2016) Tobacco Prices Going Way Up in DoD stores — You'll Pay the Same as Civilians. [online] Available at: `https://www.militarytimes.com/news/your-military/2016/04/29/tobacco-prices-going-way-up-in-dod-stores-you-ll-pay-the-same-as-civilians/` [Accessed 18 January, 2020]

[48] Kitagawa, T. (2015) A Test for Instrument Validity. *Econometrica*, 83 (5): pp. 2043–2063.

[49] Kramarow, Ellen A, and Patricia N. Pastor. (2012) The Health of Male Veterans and Nonveterans aged 25–64: United States, 2007–2010. NCHS data brief, Hyattsville, MD: National Center for Health Statistics.

[50] Kutinova, A. (2009) Paternity Deferments and the Timing of Births: U. S. Natality During the Vietnam War. *Economic Inquiry*, 47(2): pp. 351-365.

[51] Lee, D. (2009) Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies*, 76 (3): pp. 1071–1102.

[52] London, A., Herd, P., Miech, R. and Wilmoth, J. (2017). The Influence of Men's Military Service on Smoking Across the Life Course. *Journal of Drug Issues*, 47(4), pp.562-586.

[53] Lown, E., Greenfield, T., and Rogers, J. (2007). Health Effects from Drinking: Type, Severity, and Associated Drinking Patterns Based on Qualitative and Quantitative Questions in a Methodological Survey. *Substance Use & Misuse*, 42 (5), pp.793–810.

[54] Lowry, F. (2016) Unique Smoking Cessation Initiative Successful in Veterans. Medscape. [online] Available at: `https://www.medscape.com/viewarticle/873270` [Accessed 5 Mar. 2018].

[55] Manski, C. (1990) Nonparametric Bounds on Treatment Effects. *American Economic Review*, 80(2): pp. 319–323.

[56] Massoglia, M., and Pridemore, W. (2015) Incarceration and Health. *Annual Review of Sociology*, 41(1): pp. 291–310.

[57] Mourifié, I., and Wan, Y. (2017) Testing Local Average Treatment Effect Assumptions. *Review of Economics and Statistics*, 99(2), pp. 305-313.

[58] Oreopoulos, P. and Salvanes, K. (2011) Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, 25(1), pp. 159-184.

[59] Patten, S., Williams, J., Wang, J. (2006) Mental Disorders in a Population Sample with Musculoskeletal Disorders. *BMC Musculoskeletal Disorders*, 7(37), `https://doi.org/10.1186/1471-2474-7-37`

[60] Peterson, C. (1998) *Avoidance And Evasion Of Military Service: An American History, 1626-1973*. San Francisco: International Scholars Publications.

[61] Pietz, K. and Petersen, L. (2007) Comparing Self-Reported Health Status and Diagnosis-Based Risk Adjustment to Predict 1- and 2 to 5-Year Mortality. *Health Services Research*, 42(2): pp. 629-643.

[62] Roy, A. (1951) Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers* 3(2): pp. 135-146.

[63] Schmitz, L., and Conley., D. (2016) The Long-Term Consequences of Vietnam-Era Conscription and Genotype on Smoking Behavior And Health. *Behavior Genetics*, 46 (1): 43-58.

[64] Schnittker, J., and Bacak, V. (2014) The Increasing Predictive Validity of Self-Rated Health. *PLoS ONE* 9(1): e84933. Accessed on April 15, 2015. `http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0084933&type=printable.`

[65] Selective Service System. (1970) *Semiannual Report of the Director of Selective Service for the Period January 1 to June 30, 1970*, Washington, DC: U.S. Government Printing Office.

[66] ———— (1971a) *Semiannual Report of the Director of Selective Service for the Period July 1 to December 31, 1970*, Washington, DC: U.S. Government Printing Office.

[67] ———— (1971b) *Semiannual Report of the Director of Selective Service for the Period January 1 to June 30, 1971*, Washington, DC: U.S. Government Printing Office.

[68] ———— (1972a) *Semiannual Report of the Director of Selective Service for the Period July 1 to December 31, 1971*, Washington, DC: U.S. Government Printing Office.

[69] ———— (1972b) *Semiannual Report of the Director of Selective Service for the Period January 1 to June 30, 1972*, Washington, DC: U.S. Government Printing Office.

[70] ———— (1973) *Semiannual Report of the Director of Selective Service for the Period July 1 to December 31, 1972*, Washington, DC: U.S. Government Printing Office.

[71] Seltzer, C. and Jablon, S. (1974) Effects of Selection on Mortality. *American Journal of Epidemiology*, 100(5): pp.367-372.

[72] Shane, Leo III. (2019) Younger Troops Would be Barred from Buying Tobacco Products under New Plan. [online] Available at: `https://www.militarytimes.com/news/pentagon-congress/2019/05/20/younger-troops-would-be-barred-from-buying-tobacco-products-under-new-plan/` [Accessed 18 Jan. 2020]

[73] Shapiro, A. and Striker, J. (1970) *Mastering the Draft; A Comprehensive Guide for Solving Draft Problems*. Boston, MA: Little, Brown and Company.

[74] Singleton, P. (2009). The Effect of Disability Insurance on Health Investment: Evidence from the Veterans Benefits Administration's Disability Compensation Program. *Journal of Human Resources*, 44(4), pp.998-1022.

[75] Siminski, P., and Ville, S. (2011). Long-Run Mortality Effects of Vietnam-Era Army Service: Evidence From Australia's Conscription Lotteries. *American Economic Review*, 101 (3): pp. 345-349.

[76] Siminski, P., and Ville, S. (2012). I Was Only Nineteen, 45 Years Ago: What Can We Learn From Australia's Conscription Lotteries?. *Economic Record*, 88 (282): pp. 351-371.

[77] Stoye, J. (2009) More on Confidence Intervals for Partially Identified Parameters. *Econometrica*, 77(4): pp. 1299–1315.

[78] Suttler, D. (1970) IV-F; *A Guide to Medical, Psychiatric, and Moral Unfitness Standards for Military Induction*. New York, NY: Grove Press.

[79] Tarr, C. W. (1981) *By the Numbers: The Reform of the Selective Service System 1970-72*, Washington DC: National Defense University Press.

[80] Teachman, J. (2011). Are veterans healthier? Military service and health at age 40 in the all-volunteer era. *Social Science Research*, 40(1): pp. 326-335.

[81] Thompson, J., Pranger, T., VanTil, L. McColl, M., Besemann, M., Shubaly, C., Pedlar, D. (2015) Disability Correlates in Canadian Armed Forces Regular Force Veterans. *Disability and Rehabilitation*, 37(10): pp. 884-891.

[82] Torche, F. and Rauf, T. (2021) The Transition to Fatherhood and Health of Men. *Journal of Marriage and Family*, 83(2): pp. 446-465.

[83] U.S. Department of Health and Human Services. (2014) *The Health Consequences of Smoking – 50 Years of Progress, A Report of the Surgeon General*. [online] Available at: https://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf [Accessed 29 Dec. 2016].

[84] Wang, X. and Flores-Lagunes, A. (2020) Conscription and Military Service: Do They Result in Future Violent and Non-Violent Incarcerations and Recidivism? *Journal of Human Resources*, in press.

[85] Watkins, S., and Sherk, J. (2008) *Who Serves in the U.S. Military? The Demographics of Enlisted Troops and Officers*. Washington, DC: The Heritage Foundation.

[86] Zajacova, A. (2006). Education, gender, and mortality: Does schooling have the same effect on mortality for men and women in the US? *Social Science & Medicine*, 63(8), pp. 2176-2190.

[87] Zhang, J.L. and Rubin, D. (2003). Estimation of Causal Effects Via Principal Stratification When Some Outcomes are Truncated by 'Death'. *Journal of Educational and Behavioral Statistics*, 28, pp. 353–368.

[88] Zhang, J. L., Rubin, D., and Mealli, F. (2008) Evaluating the Effects of Job Training Programs on Wages through Principal Stratification. In *Advances in Econometrics*, Vol XXI, edited by D. Millimet et al. Amsterdam: North Holland.

Table 1: Summary Statistics of General Health Outcomes and Health Behaviors

	Whites			Nonwhites		
Variable	Veterans	Nonveterans	Difference	Veterans	Nonveterans	Difference
NHIS 1974-1981						
Sample size	9,257	19,824		1,139	2,883	
Draft-eligible	0.5731	0.3791	0.1940***	0.5315	0.4108	0.1207***
	[0.0057]	[0.0035]	[0.0067]	[0.0168]	[0.0093]	[0.0192]
Fair or Poor Health	0.0344	0.0337	0.0007	0.0393	0.0330	0.0063
	[0.0012]	[0.0008]	[0.0014]	[0.0052]	[0.0028]	[0.0059]
Activity Limitation	0.1399	0.1435	-0.0036	0.1380	0.1452	-0.0072
	[0.0030]	[0.0021]	[0.0037]	[0.0089]	[0.0058]	[0.0106]
Activity Unable	0.0363	0.0352	0.0010	0.0358	0.0360	-0.0002
	[0.0013]	[0.0008]	[0.0015]	[0.0049]	[0.0032]	[0.0058]
Current Smoker	0.4369	0.3059	0.1310***	0.4516	0.3074	0.1442***
	[0.0126]	[0.0085]	[0.0150]	[0.0345]	[0.0220]	[0.0407]
NHIS 1982-1996						
Sample size	14,993	32,370		2,263	6,925	
Draft-eligible	0.5608	0.3864	0.1744***	0.5399	0.4272	0.1127***
	[0.0049]	[0.0032]	[0.0058]	[0.0135]	[0.0072]	[0.0153]
Fair or Poor Health	0.0880	0.0838	0.0042	0.0779	0.0787	-0.0007
	[0.0022]	[0.0015]	[0.0026]	[0.0073]	[0.0043]	[0.0085]
Activity Limitation	0.1297	0.1186	0.0111***	0.1466	0.1097	0.0368***
	[0.0029]	[0.0019]	[0.0036]	[0.0087]	[0.0044]	[0.0098]
Activity Unable	0.0398	0.0351	0.0046**	0.0441	0.0271	0.0169**
	[0.0017]	[0.0011]	[0.0021]	[0.0064]	[0.0033]	[0.0072]
Work Limitation	0.0956	0.0846	0.0110***	0.1033	0.0775	0.0259***
	[0.0025]	[0.0017]	[0.0031]	[0.0078]	[0.0041]	[0.0088]
Work Unable	0.0462	0.0419	0.0043**	0.0492	0.0334	0.0158**
	[0.0018]	[0.0011]	[0.0021]	[0.0064]	[0.0033]	[0.0071]
Current Smoker	0.4062	0.3062	0.1000***	0.4057	0.3043	0.1014
	[0.0212]	[0.0139]	[0.0253]	[0.0535]	[0.0299]	[0.0617]
NHIS 1997-2005						
Sample size	5,472	14,235		999	3,380	
Draft-eligible	0.5575	0.3902	0.1673***	0.5199	0.4222	0.0978***
	[0.0069]	[0.0044]	[0.0083]	[0.0154]	[0.0093]	[0.0170]
Fair or Poor Health	0.1041	0.0874	0.0167**	0.1116	0.0896	0.0220
	[0.0046]	[0.0026]	[0.0054]	[0.0125]	[0.0065]	[0.0142]
Activity Limitation	0.1535	0.1100	0.0436***	0.1584	0.1070	0.0514***
	[0.0053]	[0.0028]	[0.0060]	[0.0124]	[0.0058]	[0.0136]
Work Limitation	0.1362	0.1002	0.0360***	0.1459***	0.0949	0.0510***
	[0.0048]	[0.0026]	[0.0054]	[0.0119]	[0.0055]	[0.0132]
Work Unable	0.0807	0.0613	0.0194***	0.0907	0.0560	0.0347***
	[0.0038]	[0.0021]	[0.0043]	[0.0103]	[0.0046]	[0.0113]
Current Smoker	0.2759	0.1822	0.0938***	0.2538	0.1791	0.0747***
	[0.0102]	[0.0059]	[0.0118]	[0.0216]	[0.0116]	[0.0248]
Current Drinker	0.6249	0.5915	0.0333***	0.6446	0.5669	0.0777***
	[0.0095]	[0.0061]	[0.0114]	[0.0214]	[0.0123]	[0.0251]
NHIS 2006-2013						
Sample size	3,875	9,539		819	2,596	
Draft-eligible	0.5527	0.3850	0.1677***	0.5169	0.4088	0.1081***
	[0.0084]	[0.0053]	[0.0096]	[0.0222]	[0.0102]	[0.0247]
Fair or Poor Health	0.1173	0.0891	0.0282***	0.1264	0.0879	0.0385**
	[0.0071]	[0.0040]	[0.0081]	[0.0161]	[0.0083]	[0.0181]
Activity Limitation	0.1555	0.1147	0.0407***	0.2002	0.1056	0.0947***
	[0.0075]	[0.0044]	[0.0087]	[0.0163]	[0.0084]	[0.0185]
Work Limitation	0.1397	0.1032	0.0365***	0.1731	0.0960	0.0771***
	[0.0071]	[0.0041]	[0.0083]	[0.0157]	[0.0077]	[0.0177]
Work Unable	0.0865	0.0696	0.0169**	0.1009	0.0625	0.0384**
	[0.0060]	[0.0035]	[0.0069]	[0.0138]	[0.0073]	[0.0158]
Current Smoker	0.2191	0.1514	0.0677***	0.2495	0.1435	0.1060***
	[0.0107]	[0.0063]	[0.0124]	[0.0219]	[0.0122]	[0.0252]
Current Drinker	0.6116	0.6261	-0.0145	0.6829	0.5917	0.0912***
	[0.0123]	[0.0081]	[0.0148]	[0.0219]	[0.0151]	[0.0264]

Notes: Standard errors are shown in squared brackets; * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Table 2: Relationship between latent principal strata and observed military service status (D) and eligibility to draft status (Z)

		Z	
		0	1
D	0	Never takers (*nt*) & Compliers (*c*)	Never takers (*nt*) & Defiers (*df*)
	1	Volunteers (*at*) & Defiers (*df*)	Volunteers (*at*) & Compliers (*c*)

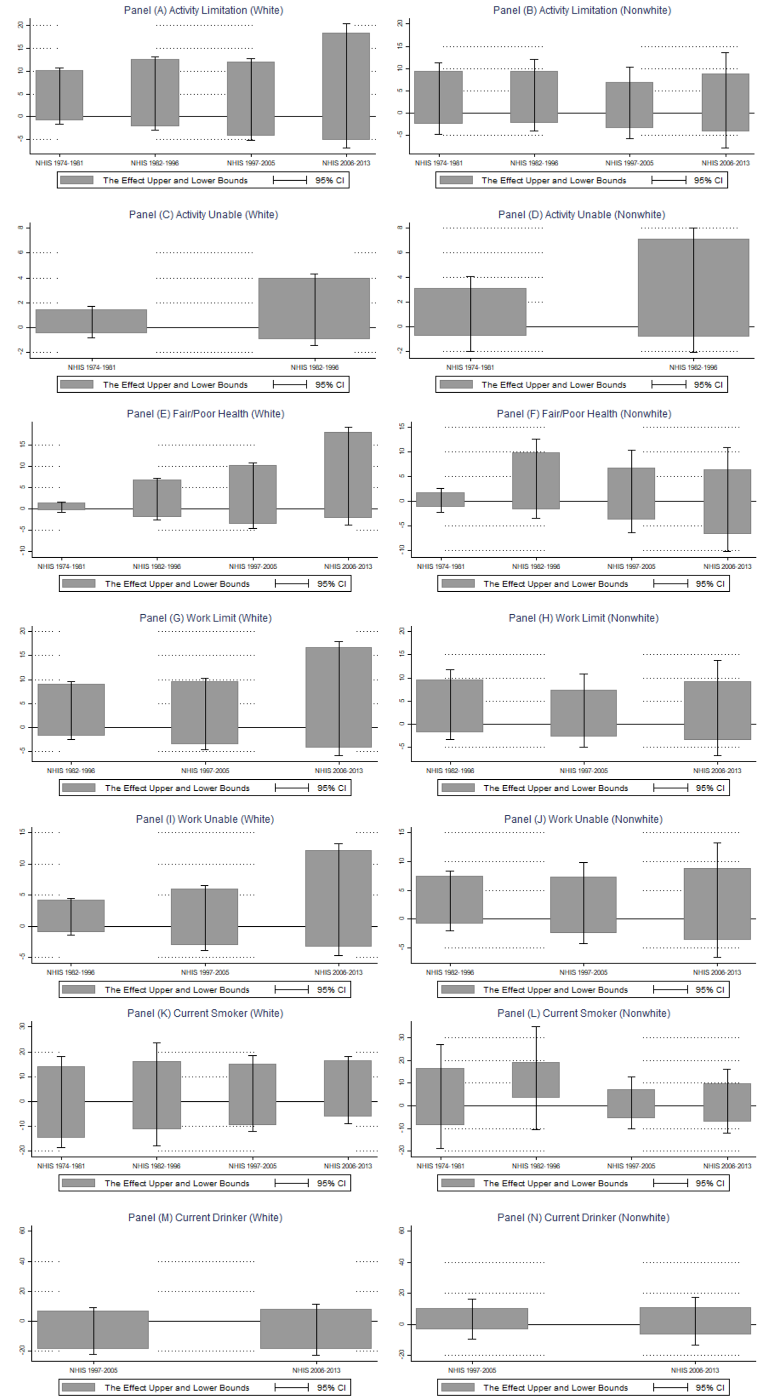
Note: In the terminology of Imbens and Angrist (1994), volunteers are always takers (*at*).

Table 3: OLS and ITT Estimates on General Health Outcomes and Health Behaviors

	Whites		Nonwhites	
NHIS 1974-1981				
Estimates	OLS	ITT	OLS	ITT
Activity Limitation	-0.0036	0.0108***	-0.0072	-0.0017
S.E.	[0.0037]	[0.0034]	[0.0106]	[0.0098]
Activity Unable	0.0010	-0.0005	-0.0002	-0.0015
S.E.	[0.0015]	[0.0014]	[0.0058]	[0.0053]
Fair and Poor Health	0.0007	-0.0011	0.0063	-0.0036
S.E.	[0.0014]	[0.0013]	[0.0059]	[0.0051]
Current Smoker	0.1310***	-0.0007	0.1442***	0.0361
S.E.	[0.0150]	[0.0142]	[0.0407]	[0.0376]
NHIS 1982-1996				
Estimates	OLS	ITT	OLS	ITT
Activity Limitation	0.0111***	0.0063*	0.0368***	0.0108
S.E.	[0.0036]	[0.0033]	[0.0098]	[0.0082]
Activity Unable	0.0046**	0.0013	0.0169**	0.0083
S.E.	[0.0021]	[0.0019]	[0.0072]	[0.0059]
Fair and Poor Health	0.0042	-0.0020	-0.0007	0.0060
S.E.	[0.0026]	[0.0024]	[0.0085]	[0.0076]
Work Limit	0.0110	0.0040	0.0259***	0.0057
S.E.	[0.0031]	[0.0028]	[0.0088]	[0.0073]
Work Unable	0.0043**	0.0011	0.0158**	0.0083
S.E.	[0.0021]	[0.0019]	[0.0071]	[0.0059]
Current Smoker	0.1000***	-0.0036	0.1014	0.0248
S.E.	[0.0253]	[0.0237]	[0.0617]	[0.0539]
NHIS 1997-2005				
Estimates	OLS	ITT	OLS	ITT
Activity Limitation	0.0436***	-0.0103**	0.0514***	0.0013
S.E.	[0.0060]	[0.0051]	[0.0136]	[0.0112]
Fair and Poor Health	0.0167***	-0.0128***	0.0220	0.0021
S.E.	[0.0054]	[0.0045]	[0.0142]	[0.0117]
Work Limit	0.0360***	-0.0111**	0.0510***	0.0037
S.E.	[0.0053]	[0.0046]	[0.0132]	[0.0104]
Work Unable	0.0194***	-0.0110***	0.0347***	-0.0027
S.E.	[0.0043]	[0.0037]	[0.0113]	[0.0087]
Current Smoker	0.0938***	-0.0088	0.0747***	-0.0162
S.E.	[0.0118]	[0.0105]	[0.0247]	[0.0202]
Current Drinker	0.0333***	-0.0058	0.0777***	0.0616***
S.E.	[0.0114]	[0.0105]	[0.0251]	[0.0217]
NHIS 2006-2013				
Estimates	OLS	ITT	OLS	ITT
Activity Limitation	0.0407***	-0.0025	0.0947	-0.0053
S.E.	[0.0087]	[0.0077]	[0.0185]	[0.0156]
Fair and Poor Health	0.0282***	0.0132	0.0385**	-0.0227
S.E.	[0.0081]	[0.0071]	[0.0181]	[0.0149]
Work Limit	0.0365***	-0.0016	0.0771***	-0.0032
S.E.	[0.0083]	[0.0072]	[0.0177]	[0.0150]
Work Unable	0.0169**	-0.0022	0.0384	-0.0110
S.E.	[0.0069]	[0.0063]	[0.0158]	[0.0132]
Current Smoker	0.0677***	-0.0066	0.1060***	-0.0442**
S.E.	[0.0124]	[0.0112]	[0.0252]	[0.0215]
Current Drinker	-0.0145	0.0004	0.0912***	0.0245
S.E.	[0.0148]	[0.0138]	[0.0264]	[0.0254]

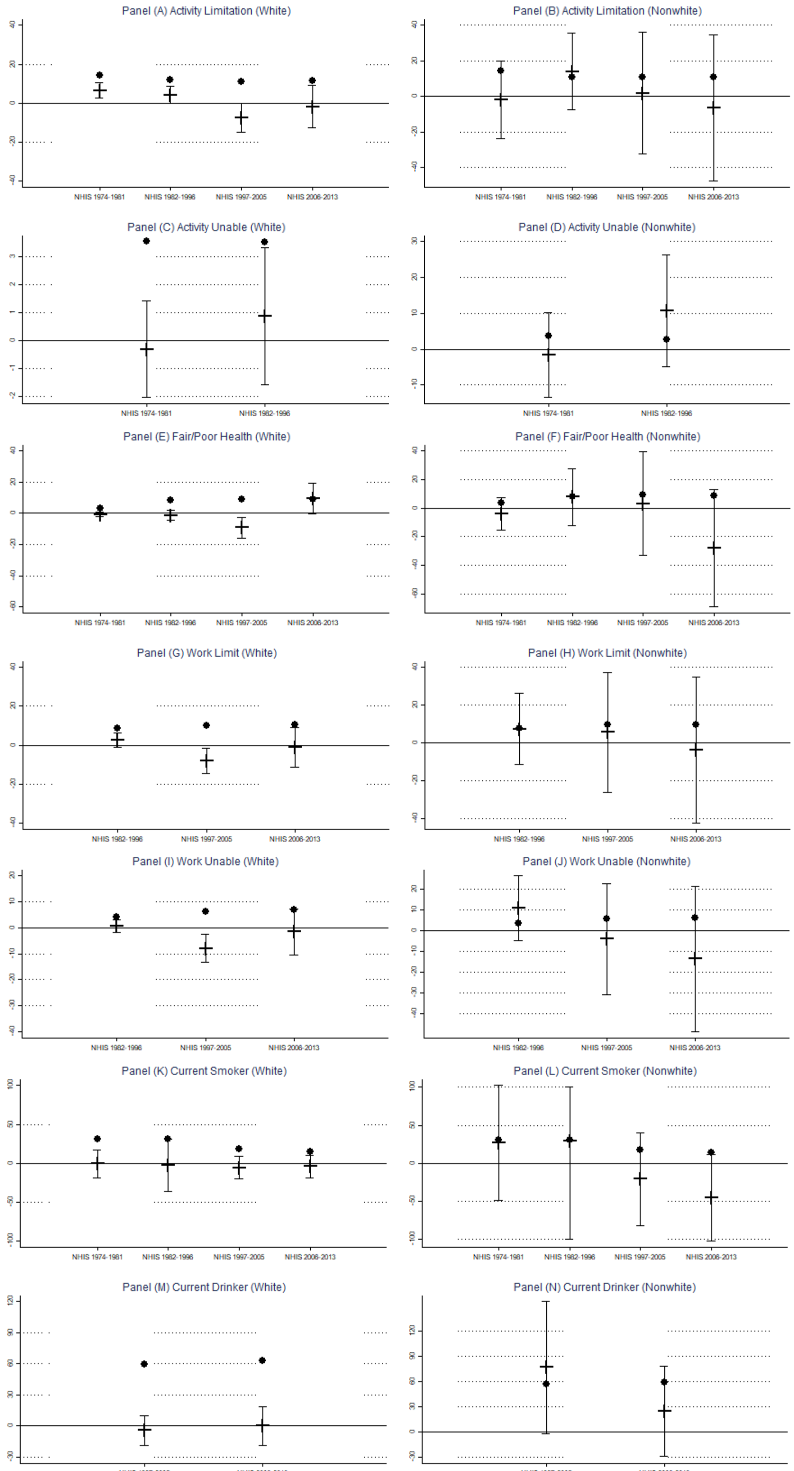
Notes: OLS (respectively, ITT) estimates represent the difference in average outcomes between veterans and nonveteans (eligible to draft and ineligible to draft). Standard errors of estimates are shown in squared brackets. * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Figure 1. Est. Bounds on the Net Effect of the Draft Lotteries on General Health Outcomes and Behaviors of Never Takers



Notes: Panels in the left column correspond to whites while panels in the right column correspond to nonwhites. Within each column, each panel corresponds to a general health outcome or health behavior. Within each panel, results are presented for each of the survey periods for which the general health outcome or behavior is available.

Figure 2. Estimated Local Effect of Military Service on General Health Outcomes and Health Behaviors of Complier Veterans



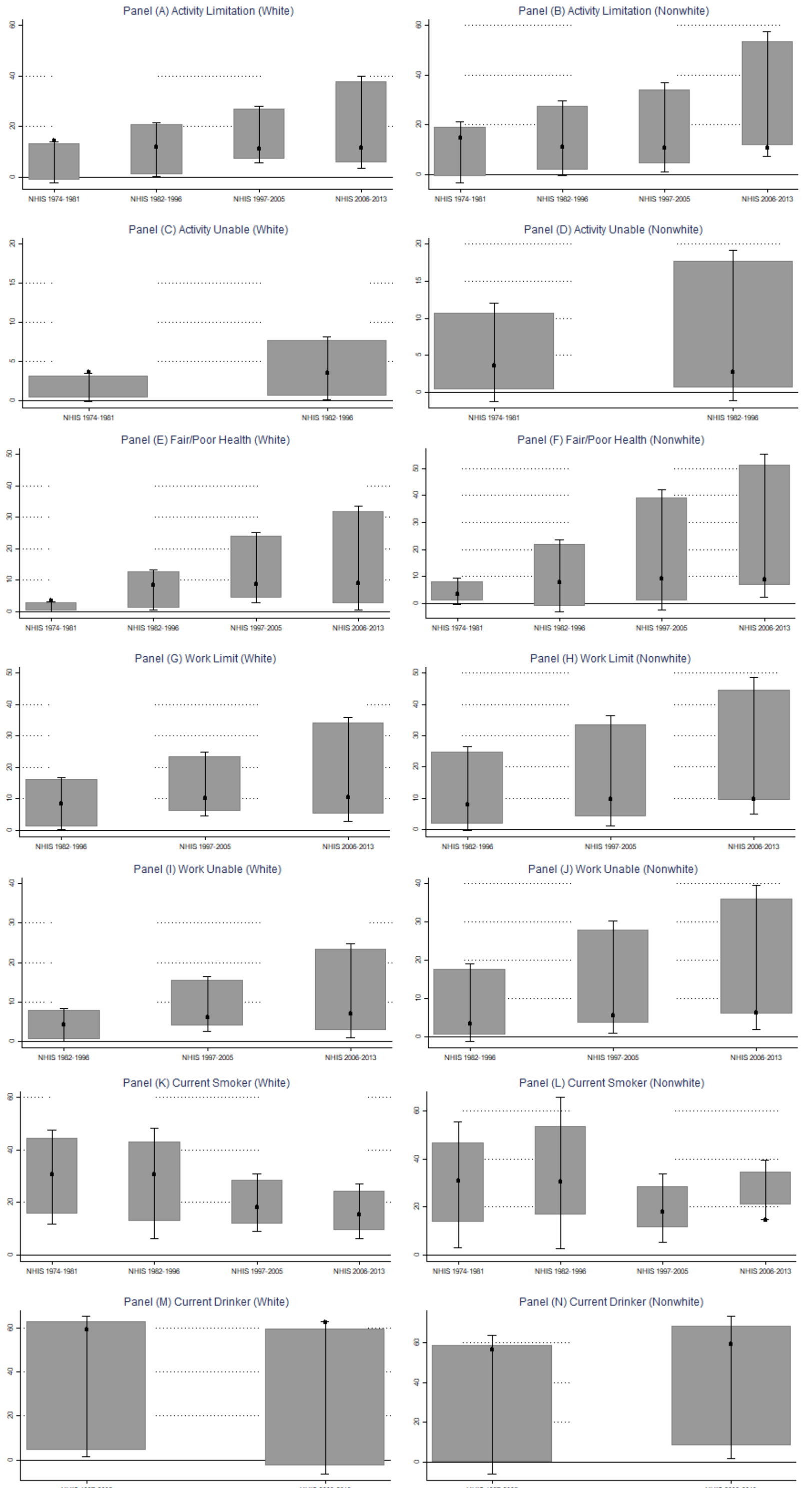
Notes: Panels in the left column correspond to whites while panels in the right column correspond to nonwhites. Within each column, each panel corresponds to a general health outcome or health behavior. Within each panel, results are presented for each of the survey periods for which the general health outcome or behavior is available. For each outcome, the point estimate is represented by a cross, the 95% confidence interval is represented by a vertical line, and the estimated non-veteran mean ($E[Y|D = 0]$) is represented by a dot.

Table 4: Analysis of Two Pre-Draft Characteristics for the Latent Subpopulations

Characteristics	Never Takers (nt)	Volunteers (at)	Compliers (c)	Differences at-nt	nt-c	at-c
NHIS 1974-1981: Whites						
Activity limitations before 1965	0.0222	0.0064	0.0068	-0.0158***	0.0154***	-0.0004
	[0.0012]	[0.0014]	[0.0031]	[0.0017]	[0.0036]	[0.0040]
High School Incompletion	0.1548	0.1191	0.0747	-0.0356***	0.0800***	0.0444**
	[0.0033]	[0.0052]	[0.0121]	[0.0057]	[0.0139]	[0.0152]
NHIS 1974-1981: Nonwhites						
Activity limitations before 1965	0.0199	0.0023	0.0083	-0.0176***	0.0116	-0.0060
	[0.0027]	[0.0025]	[0.0118]	[0.0034]	[0.0129]	[0.0133]
High School Incompletion	0.2746	0.1297	0.0114	-0.1449***	0.2632***	0.1183
	[0.0088]	[0.0155]	[0.0626]	[0.0169]	[0.0650]	[0.0726]
NHIS 1982-1996: Whites						
High School Incompletion	0.1357	0.0819	0.0334	-0.0538***	0.1024***	0.0486***
	[0.0022]	[0.0034]	[0.0086]	[0.0034]	[0.0097]	[0.0110]
NHIS 1982-1996: Nonwhites						
High School Incompletion	0.2532	0.0795	0.0712	-0.1737***	0.1820***	0.0084
	[0.0057]	[0.0097]	[0.0362]	[0.0099]	[0.0385]	[0.0430]

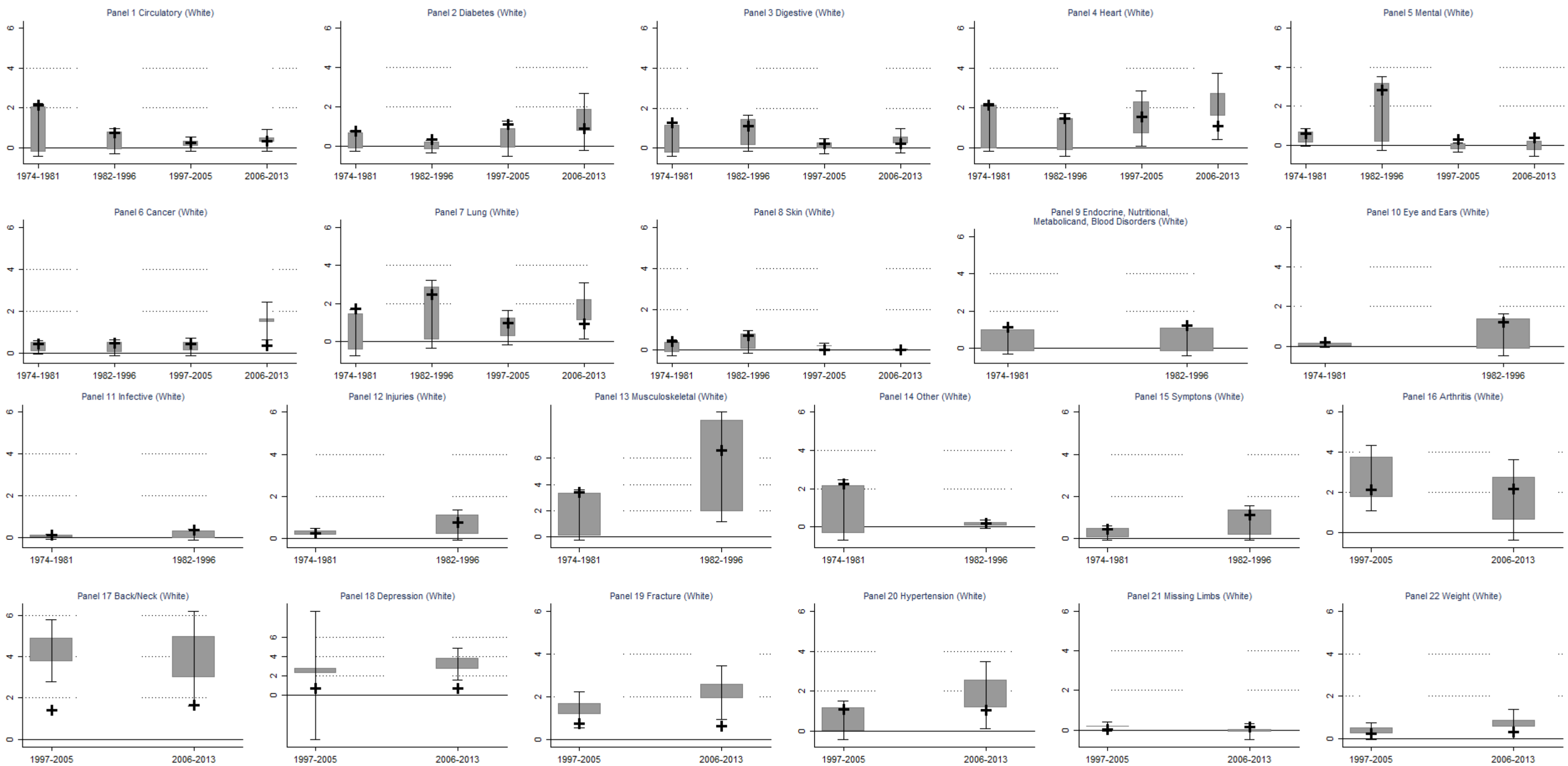
Notes: Standard errors are shown in squared brackets. The estimated differences in the last three columns indicate: * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Figure 3. Est. Bounds on the Local Effect of Military Service on General Health Outcomes and Behaviors of Volunteer Veterans



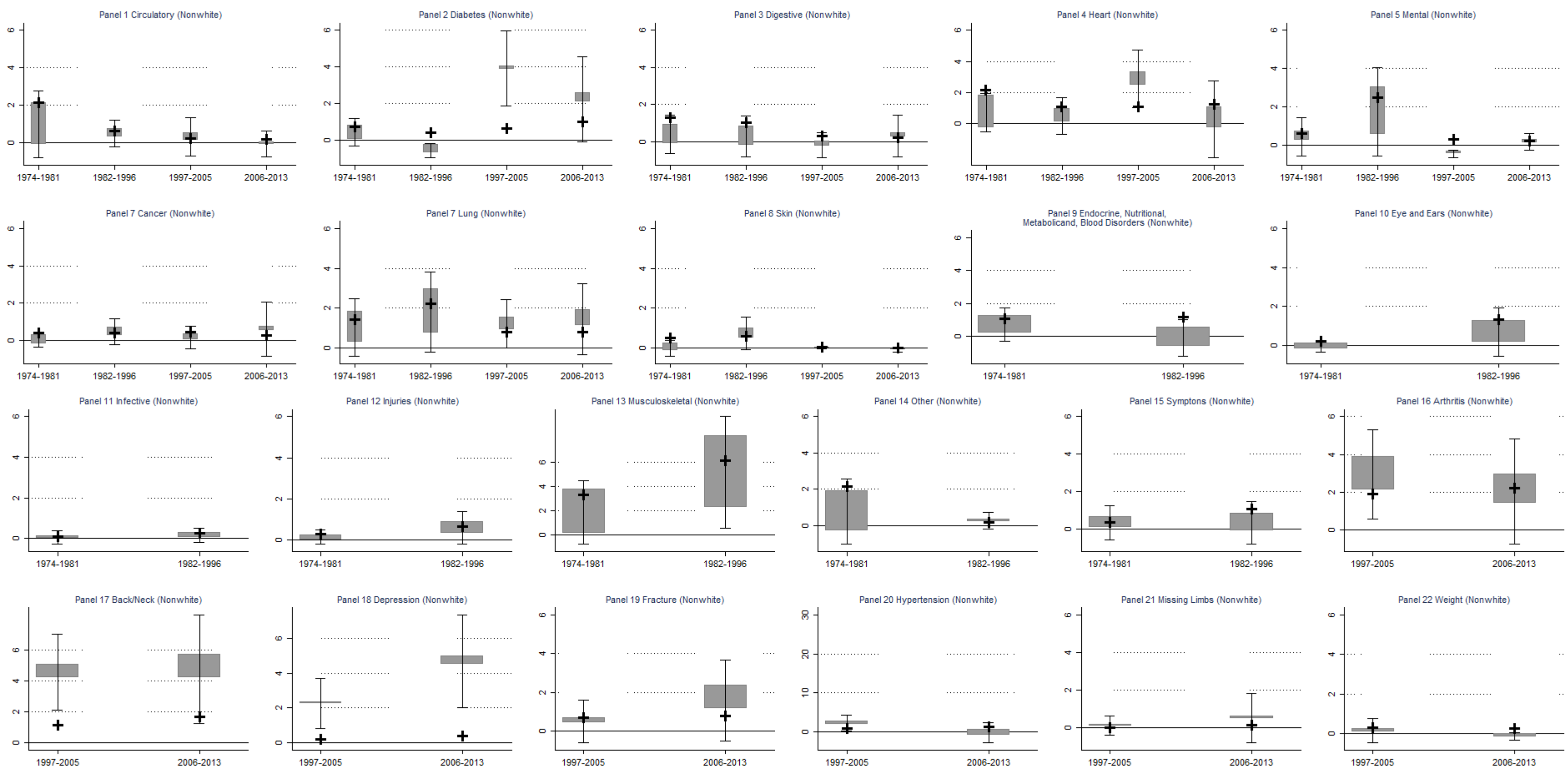
Notes: Panels in the left column correspond to whites while panels in the right column correspond to nonwhites. Within each column, each panel corresponds to a general health outcome or health behavior. Within each panel, results are presented for each of the survey periods for which the general health outcome or behavior is available. For each outcome, the estimated bounds are represented by shaded rectangles, the 95% confidence interval is represented by a vertical line, and the estimated non-veteran mean ($E[Y|D=0]$) is represented by a dot.

Figure 4. Est. Bounds on the Local Effect of Military Service on Activity-Limiting Chronic Conditions of White Volunteer Veterans



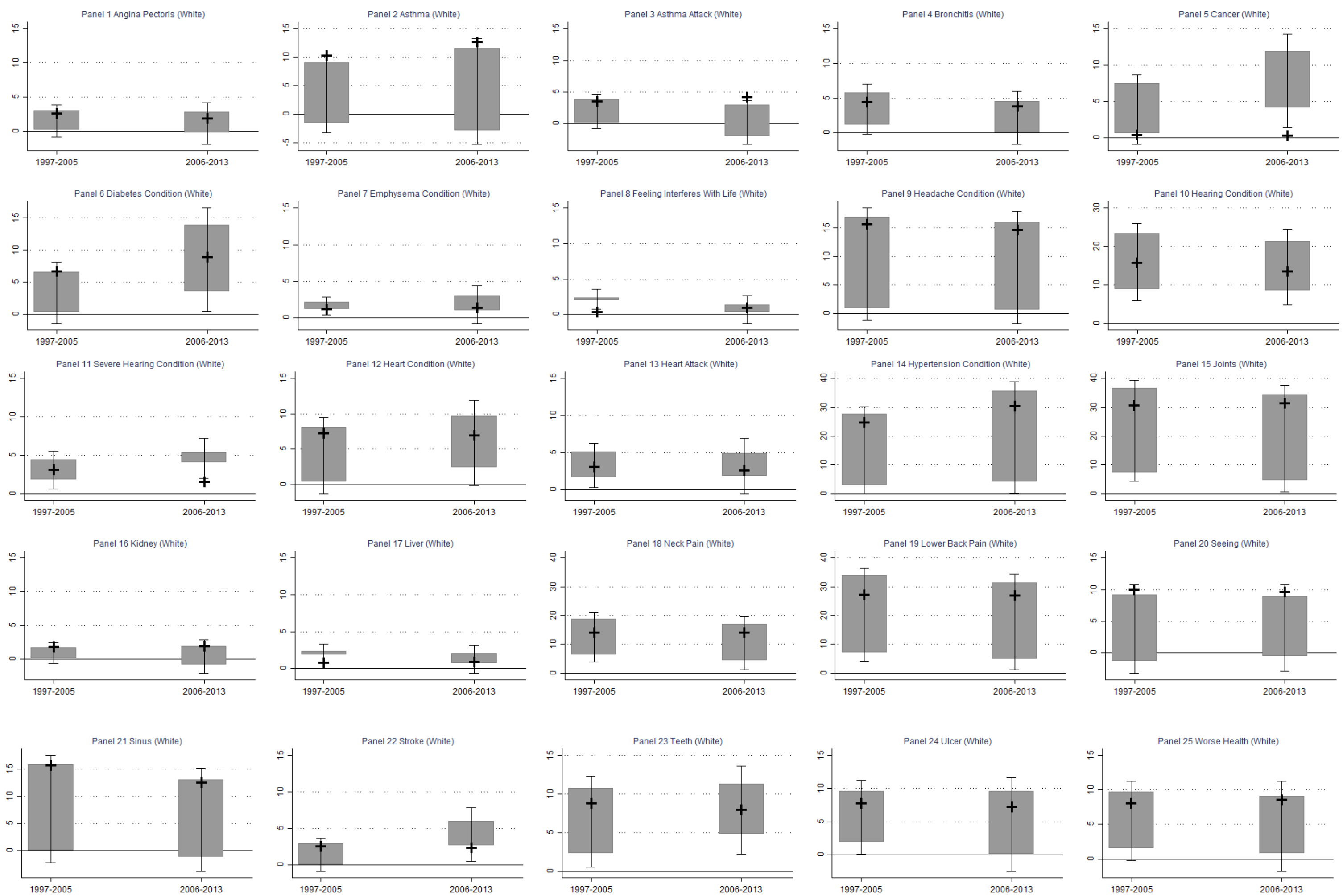
Notes: Each panel corresponds to a type of activity-limiting chronic conditions. Within each panel, results are presented for each of the survey periods for which the outcome is available. For each outcome, the estimated bounds are represented by the shaded rectangles, the 95% confidence intervals by vertical lines, and the non-veteran average outcome by a bold cross.

Figure 5. Est. Bounds on the Local Effect of Military Serv. on Activity-Limiting Chronic Conditions of Nonwhite Volunteer Veterans



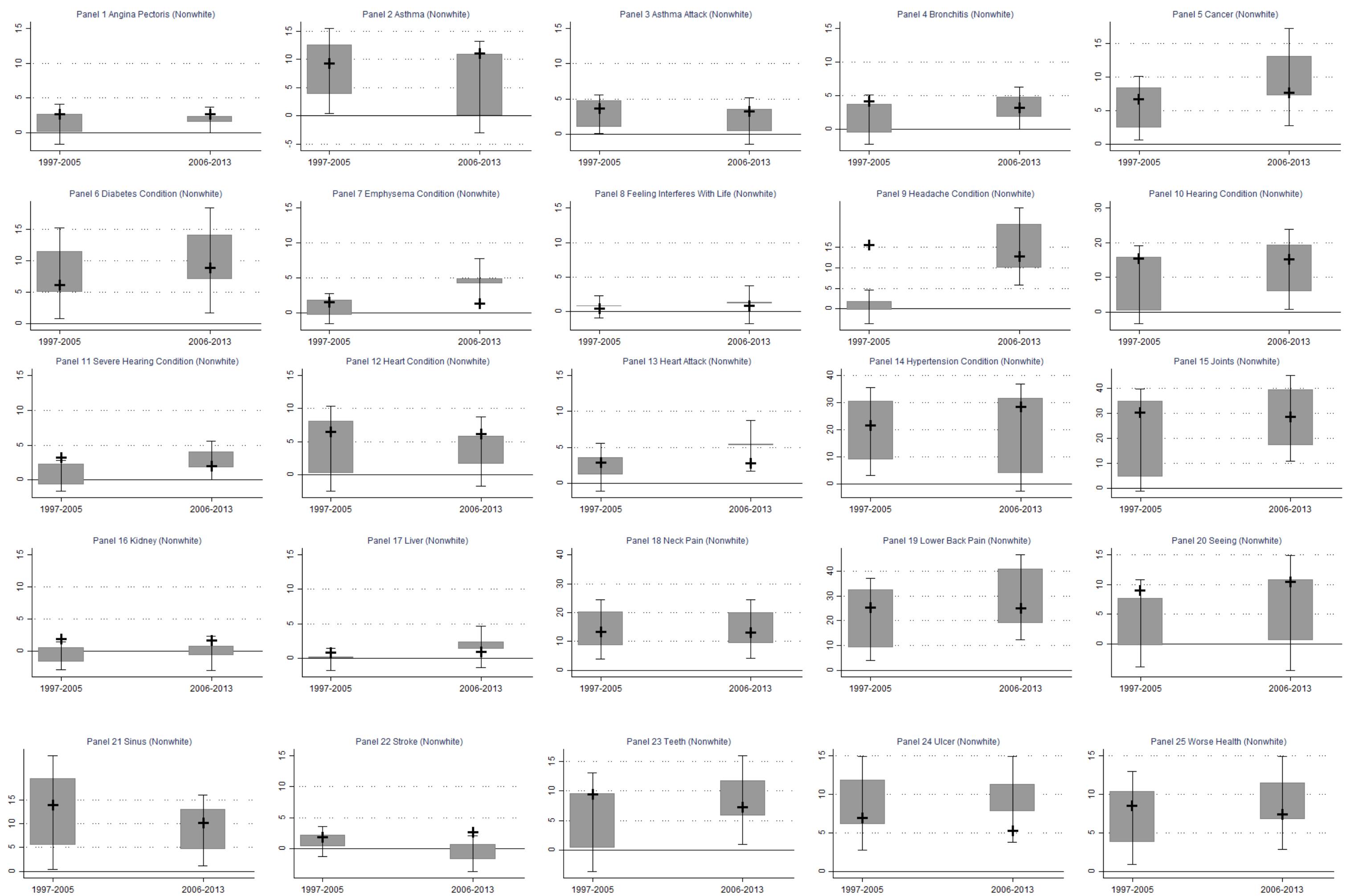
Notes: Each panel corresponds to a type of activity-limiting chronic conditions. Within each panel, results are presented for each of the survey periods for which the outcome is available. For each outcome, the estimated bounds are represented by the shaded rectangles, the 95% confidence intervals by vertical lines, and the non-veteran average outcome by a bold cross.

Figure 6. Est. Bounds on the Local Effect of Military Service on Other Chronic Conditions of White Volunteer Veterans



Notes: Each panel corresponds to a different type of chronic conditions. Within each panel, results are presented for each of the two survey periods for which the outcome is available. For each outcome, the estimated bounds are represented by the shaded rectangles, the 95% confidence intervals by vertical lines, and the non-veteran average outcome by a bold cross.

Figure 7. Est. Bounds on the Local Effect of Military Service on Other Chronic Conditions of Nonwhite Volunteer Veterans



Notes: Each panel corresponds to a different type of chronic conditions. Within each panel, results are presented for each of the two survey periods for which the outcome is available. For each outcome, the estimated bounds are represented by the shaded rectangles, the 95% confidence intervals by vertical lines, and the non-veteran average outcome by a bold cross.

Table 5: Estimates on the Military Service Effect on Mortality by December 31, 2011

	Whites			Nonwhites		
Survey Year	NHIS 1985-1996	NHIS 1997-2005	NHIS 2006-2009	NHIS 1985-1996	NHIS 1997-2005	NHIS 2006-2009
Sample Size	31795	19555	5821	6506	4324	1444
ITT	-0.0037	-0.0000	-0.0069	-0.0037	-0.0015	0.0046
S.E.	[0.0037]	[0.0039]	[0.0058]	[0.0102]	[0.0105]	[0.0131]
OLS	0.0122***	0.0055	-0.0009	0.0009	0.0068	-0.0142
S.E.	[0.0041]	[0.0045]	[0.0065]	[0.0120]	[0.0126]	[0.0129]
2SLS	-0.0246	-0.0003	-0.0558	-0.0451	-0.0220	0.0697
S.E.	[0.0249]	[0.0282]	[0.0492]	[0.1285]	[0.1636]	[0.8560]
Volunteer Veterans	(0.0221, 0.1141)	(0.0107, 0.0908)	(0.0092, 0.0315)	(0.0157, 0.1090)	(0.0114, 0.0895)	(-0.0274, 0.0007)
95% CIs	[0.0117, 0.1228]	[-0.0008, 0.1006]	[-0.0072, 0.0457]	[-0.0140, 0.1350]	[-0.0178, 0.1147]	[-0.0525, 0.0161]
All Veterans	(0.0124, 0.0853)	(0.0084, 0.0715)	(-0.0026, 0.0156)	(0.0065, 0.0858)	(0.0068, 0.0743)	(-0.0165, 0.0085)
95% CIs	[0.0006, 0.0964]	[-0.0053, 0.0843]	[-0.0239, 0.0342]	[-0.0329, 0.1241]	[-0.0353, 0.1151]	[-0.0619, 0.0486]

Notes: Standard errors shown in brackets are based on 5000 rounds of bootstrap. For the estimated bounds for Volunteer Veterans and All Veterans, Imbens and Manski (2004) 95% CIs for the parameter of interest are shown in brackets. For point estimates, * significant at 10% level; ** significant at 5% level; *** significant at 1% level.