The Causal Effect of Military Conscription on Crime and the Labor Market*

Randi Hjalmarsson†
University of Gothenburg and CEPR

Matthew J. Lindquist††
SOFI, Stockholm University

August 30, 2016

Abstract

This paper uses detailed individual register data to identify the causal effect of mandatory peacetime military conscription in Sweden on the lives of young men born in the 1970s and 80s. Because draftees are positively selected into service based on their draft board test performance, our primary identification strategy uses the random assignment of potential conscripts to draft board officiators who have relatively high or low tendencies to place draftees into service in an instrumental variable framework. We find that military service significantly increases post-service crime (overall and across multiple crime categories) between ages 23 and 30. These results are driven primarily by young men with pre-service criminal histories and who come from low socioeconomic status households. Though we find evidence of an incapacitation effect concurrent with conscription, it is unfortunately not enough to break a cycle of crime that has already begun prior to service. Analyses of labor market outcomes tell similar post-service stories: individuals from disadvantaged backgrounds have significantly lower income, and are more likely to receive unemployment and welfare benefits, as a result of service, while service improves or does not affect outcomes for those at the other end of the distribution. Finally, we provide suggestive evidence that peer effects may play an important role in explaining the unintended negative impacts of military service.

Keywords: Conscription, Crime, Criminal Behavior, Draft, Military Conscription, Military Draft, Incapacitation, Labor Market, Unemployment.

JEL: H56, J08, K42.

---

* We would like to thank seminar participants at the Swedish Institute for Social Research and the Tinbergen Institute for their helpful comments and suggestions. Hjalmarsson would also like to gratefully acknowledge funding support from Vetenskapsrådet (The Swedish Research Council), Grants for Distinguished Young Researchers.
† University of Gothenburg, Department of Economics, Vasagatan 1, SE 405 30, Gothenburg, Sweden; randi.hjalmarsson@economics.gu.se
†† Swedish Institute for Social Research (SOFI), Stockholm University, Universitetsvägen 10F, 106 91 Stockholm, Sweden; matthew.lindquist@sofi.su.se.

1. Introduction

Young men in more than 60 countries around the world still face the prospect of mandatory military conscription today.[1] This occurs at a critical juncture in a young adult's life – when he is at the peak of the age-crime profile, making decisions about higher education, and entering the labor market. It is thus not surprising that conscription remains a hotly debated topic; in fact, a number of European countries have recently abolished it (France, 1996; Italy, 2005; Sweden, 2010; and Germany, 2011) while others have had failed referendums (Austria and Switzerland in 2013).[2] Yet, despite a growing body of academic literature, there is little consensus about the impact of this potentially life transforming event.

The current paper contributes to this debate by utilizing individual administrative records and a quasi-experimental research design to identify the causal impact of mandatory military conscription in Sweden on crime, both concurrent with (incapacitation) and after conscription. We complement this analysis by applying the same research design to legitimate labor market outcomes, including education, income, and welfare and unemployment benefits, as well as work-related health outcomes (sick days and disability benefits).

Mandatory military conscription in Sweden dates back to 1901 and was abolished in 2010, after a gradual decline that began upon the end of the Cold War. For most of this period, all Swedish male citizens underwent an intensive drafting procedure at age 18, including tests of physical and mental ability. These test results were reviewed by a randomly assigned officiator, who determined whether the draftee would be enlisted. It is this exogenous variation in the likelihood of serving that our analysis utilizes in an instrumental variable framework to identify the causal effect of conscription on crime. In addition, for a subset of cohorts for whom

[1] See the CIA's World Factbook (https://www.cia.gov/library/publications/the-world-factbook/fields/2024.html) and http://chartsbin.com/view/1887 for a summary of this data.

[2] Though the U.S. moved to an all-volunteer military in 1973, young men ages 18 to 26 are still required to register for the draft. Today, the US is debating extending this requirement to young women. http://www.nbcnews.com/news/us-news/military-officials-women-should-register-draft-n509851

we know the exact dates of service, we utilize a matching framework to identify the incapacitation effects of service.

There are a number of channels through which military conscription may affect both contemporaneous and future criminal behavior. Conscription may decrease contemporaneous criminal behavior by incapacitating young men, i.e. keeping them otherwise engaged and isolated from society. On the other hand, conscripts are not under 24-hour supervision and can still commit crimes ‘after hours’ and experience increased social interactions, which could result in an increased propensity to commit highly ‘social’ crimes.[3] If conscription does “incapacitate” potential criminals, then this could reduce post-service crime by putting conscripts on a new path of lower criminal intensity. Alternatively, the promotion of democratic values and the obedience and discipline training that one receives may decrease post-service crime by helping focus young men at this high risk age. Others argue that exposure to weapons and desensitization to violence, especially during wartime, may exacerbate one’s criminal tendencies (Grossman, 1995). Military conscription may also positively or negatively impact crime through its impact on education and the labor market (Becker, 1968). Crime could decrease if conscription extends a conscripts’ social networks, is viewed as a positive signal of quality by employers, or improves his marketable skills (e.g. training as mechanics, cooks and medics), health, or physical fitness. However, crime may increase if conscription interrupts a continuous educational path, delays entry into the labor market, and reduces future labor market opportunities. Finally, exposure to a new peer group may have either positive or negative effects on criminal behavior, depending on the relative characteristics of the new and old peer groups.

The existing research yields mixed results, with respect to both labor market and crime outcomes. Angrist’s (1990) seminal study found that Vietnam draftees in the U.S. had lower

[3] This parallels the school crime literature, where Jacob and Lefgren (2003) and Luallen (2006) have found an incapacitation effect of schooling on property crime but an exacerbating effect on violent crime, which they argue results from increased social interactions.

earnings than non-draftees; subsequent papers (Angrist and Chen, 2011; Angrist, Chen, and Song, 2011) find that this gap closes over time, such that by age 50, draftees are on par with non-draftees.[4] With respect to peacetime service, Imbens and van der Klaauw (1995) find lower wages for Dutch veterans, Grenet et al (2011) and Bauer et al (2009) find no impact on wages for British and German cohorts coming of age just after the abolition of conscription, and Card and Cardoso (2012) find a small positive effect on earnings for low-educated men in Portugal.[5] Bingley et al. (2014) find large earnings losses for high ability men in Denmark. Hanes et al. (2010) and Grönqvist and Lindqvist (2016) find a positive effect of conscription on Swedish labor market outcomes, but only the latter, which focuses on officer training, seriously addresses the biases arising from the endogenous selection process.[6, 7]

Few papers study the effect of conscription on crime in a quasi-experimental setting.[8] Those studying Vietnam Veterans in the U.S. find some evidence that conscription causes an increase in violent crimes (Rohlfs, 2010 and Lindo and Stoecker, 2012), while Siminski et al (2016) do not find an effect on violent crime in Australia during Vietnam era conscription. Just two papers consider the effects of peacetime conscription. In Argentina, where males are randomly assigned eligibility based on the last three digits of their national identity number, Galiani et al (2011) find that conscription increases crime, especially property and white collar crimes, and decreases labor market outcomes; these effects are even larger for wartime draftees. Finally, for a subset of the 1964 Danish birth cohort, Albaek et al. (forthcoming) find that

[4] Likewise, Siminski's (2013) study of Australian Vietnam draftees finds a negative employment effect.
[5] Maurin and Xenogiani (2007) use the abolition of conscription in France to study the effect of schooling on wages.
[6] Using a regression discontinuity design, Grönqvist and Lindqvist (2016) find that officer training in the military significantly increases the probability of becoming a civilian manager.
[7] Albrecht et al. (1999) also report regression coefficients that (in some specifications) show a positive return to military service in Sweden. But as their paper is about explaining the negative returns to time spent out of work (due to, e.g., maternity leave) and not about the effects of military service on earnings, they do not comment on these coefficients nor do they specifically address the issue of selection into military service.
[8] Beckerman and Fontana's (1989) survey of the early criminology and psychology literature finds that Vietnam veterans do not have higher arrest rates than non-veterans. A handful of studies find a positive effect of being a Vietnam Vet on violent crime, but these are often restricted to those in combat or individuals with mental health problems. See Yager et al (1984), Resnick et al (1989), and Yesavage (1983).

service reduces property crime among men with previous convictions for up to four years (starting from the year in which they begin military service). However, their data do not allow them to cleanly estimate an incapacitation effect separately from a post-service effect.

What can explain these diverse findings? First, the effect of conscription may change over the lifecycle. For an outcome like crime, which peaks as a young adult, focusing on crime after age 40, as done in some previous studies (Siminski et al, 2013; Galiani et al, 2011), may skew the results. Second, the conscription ‘experience’ varies greatly across studies. While peacetime versus wartime conscription is the most obvious example, other differences may emerge as countries approach the end of their mandatory conscription regimes. Third, measured differences may be related to differences in how the causal effect is identified. Because of the selection process involved in military service, one cannot simply compare outcomes for those who do and do not serve. While the most convincing studies rely on random assignment to service (i.e. lotteries), one should also be interested in the effect of service in countries without a lottery to use as an identification strategy. Several studies do this by comparing cohorts before and after the abolition of conscription, which can yield different results than the lottery design since: (i) the conscription experience likely differs when it is about to be abolished, (ii) it may include general equilibrium effects, and (iii) the average and marginal individuals ‘treated’ may not be comparable across studies. If conscription has heterogeneous effects, then it is not surprising if studies with different identification strategies find different effects.

The current paper addresses a number of limitations of the existing research. First, we look at the effect of conscription on more modern cohorts who come of age in the 1990s rather than during the Vietnam War or soon after.[9] Second, detailed individual register data allow us

[9] Anderson and Rees (2015) also study more modern data – namely those deployed in the Iraq war from Fort Carson, Colorado from 2001 to 2009; they conclude that never-deployed units have a greater impact on crime and public safety than units recently returned from combat. Bingley et al. (2014) also study more recent cohorts of Danish men born 1976-1983, with an emphasis on the labor market. They find a large negative effect on earnings, but report a zero effect of service on non-vehicular crime at the extensive margin for men aged 26-35 at all quartiles of the ability distribution.

to study crimes committed when the men are young adults (i.e. before the age of 30 for all cohorts). Third, the Swedish register data allow for one of the most comprehensive analyses to date of the impact of peacetime service on a wide range of labor market and health outcomes, in addition to crime. Methodologically, we apply a new research design to identify the causal effects of service on crime. Finally, and in contrast to all of the existing literature, information on the exact dates of service allows us to directly estimate the incapacitation effects of service.

We use a sample from Sweden's Multigenerational Register (about 70% of the population), which is matched to longitudinal administrative income, education, tax, geographical, criminal conviction, and draft board records. Most importantly, we can identify the officiators who reviewed the draft board test results and placed the draftees into service the 1990 to 1996 test years, while we know exact dates of service for the 1997 to 2001 test cohorts.

These key data features guide the two stages of our empirical analysis. The first stage aims to cleanly identify the causal effect of service on post-service crime. To deal with the 'selection' of conscripts into service on the basis of their draft board test scores, we instrument for service with whether the draftee is assigned an officiator with a high annual service rate (that is, an officiator whose annual share of testees who serve is greater than the national share who serve in that year). We argue that this instrument is both valid and relevant when conditioning on county by test year fixed effects. With regards to the former, we provide both anecdotal and empirical evidence of random assignment to officiators; we also demonstrate random assignment using a falsification test for a sample who meets an officiator but is excluded from service for health reasons. With regards to the first stage, we find that assignment to a high service rate officiator increases the chance of service by almost eight percentage points, with an appropriately high first-stage F-statistic; similar first stage relationships are seen for heterogeneous subsamples (by criminal history and paternal education). Finally, we demonstrate that while the assignment of a high service officiator affects the chance of service,

it does not affect the type of service, including the branch (army, navy, air force), rank, or whether assigned a combat position. Hence, our analysis emphasizes the causal effect of military service rather than the reduced form effect of assignment to a high service officiator.

The baseline results are striking: military service significantly increases both the likelihood of crime and the number of crimes between ages 23 and 30 (a cleanly defined post-service period). Such a positive effect is actually seen across all crime categories, with especially significant effects for violent crime, theft, and drug and alcohol offenses. The estimated effects are quite large, oftentimes more than twice the mean of the dependent variable, and appear to be driven by those from disadvantaged backgrounds with respect to criminal history or father's education. This heterogeneous impact is also seen with respect to labor market outcomes. Individuals from disadvantaged backgrounds tend to be worse off in terms of income, unemployment and welfare while those at the upper end of the distribution have significantly higher income. There is some evidence, however, that service decreases disability benefits and the number of sick days for the whole sample.

To isolate incapacitation in the IV framework, we consider crimes committed at ages 19 and 20 for those who tested at 18. Though these results suggest an incapacitating effect of service, it cannot be ruled out the crime outcomes include some pre- or post-service crimes. Thus, the second stage of the empirical analysis matches each individual in the 1997 to 2001 test cohorts who serves in the military to one specific control individual who does not serve. We then use the exact dates of service for the treated individual to construct the counterfactual time of incapacitation for the control group and apply a matching estimator to identify the incapacitation effects of military service. We find large incapacitation effects across all crime types. Taken together, our matching and instrumental variable estimates imply that there are significant incapacitation effects of military service. Unfortunately, our analysis also suggests

that these effects are not large enough to break a cycle of crime that has already begun prior to service.

Finally, we demonstrate that individuals with criminal histories prior to service and from low socioeconomic status households are likely to be concentrated together when conscripted, leading to potentially intense social interactions. We find a strong relationship between peer criminal history prior to service and an individual’s post service crime, for just those individuals from disadvantaged backgrounds. As such, negative peer interactions appear to be one feasible explanation for the unintended negative consequences of military service. Another possible mechanism is that low skilled men are harmed by their delayed entry into the labor market during years when unemployment among young adults is unusually high.

The remainder of the paper proceeds as follows. Section 2 provides institutional details about Swedish military conscription and an overview of the data. Section 3 presents the high service rate officiator instrumental variable strategy to identify the post-service effect of conscription. Section 4 presents the instrumental variable results for crime and non-crime outcomes as well as heterogeneity and sensitivity analyses. Section 5 presents the matching framework to isolate incapacitation and results. Section 6 discusses the potential mechanisms that may explain the large effects of service on crime, highlighting the possibility of negative peer effects. Section 7 concludes.

2. Mandatory Conscription in Sweden

2.1. A Brief History

Mandatory military service in Sweden dates back to 1901. Shortly after turning 18, all Swedish male citizens underwent an intensive drafting procedure, including tests of cognitive ability, endurance, strength, and physical and mental health, the results of which determined whether one would be conscripted and the assigned unit and rank. Most individuals enlisted at age 19

or 20, for 7 to 15 months, depending on unit and rank.[10] Individuals were trained in three stages: soldiering skills, skills specific to each line of service (army, navy, air force, coastal artillery), and joint training exercises to prepare for wartime deployment.

Though peacetime conscription was officially suspended on July 1, 2010, the number of men placed into military service actually started decreasing upon the end of the Cold War and accelerating after the fall of the Berlin Wall in 1989. This decrease was formalized in the Defense Proposition of 1992. Figure 1 shows the share of men born in Sweden (by birth cohort) who were called and tested by *Värnpliktsverket* (The Swedish Conscription Authority), as well as the share who were deemed fit to serve and placed in service categories and the share who actually served in the military. Roughly 95% of all men born before 1979 were called and tested; the remaining 5% were typically non-citizens or individuals with severe mental or physical problems that exempted them from military service. At that time, noncompliance was punishable by jail and this law was rigorously enforced.

In 1995, *Värnpliktsverket* and *Vapenfristyrelsen* (The Civil Conscription Committee) were merged into a single conscription authority called *Pliktverket*. Between 1995 and 2007, the number of young males called and tested by *Pliktverket* fell by 10 percentage points, due to a more thorough pre-test screening process meant to avoid testing those most likely to be exempted for mental and physical health reasons.[11] In 2007, *Pliktverket* began using an online tool to further pre-screen potential conscripts – only a small share of the most suitable, willing and able were called for testing, and even fewer were placed in service. On July 1, 2010, Sweden adopted an all voluntary military service, though male citizens must still register with the official recruitment office (*Rekryteringsmyndigheten*) at 18. One of the most significant policy changes in recent Swedish history, its consequences have not yet been thoroughly evaluated.

---

[10] Military training was typically 7 to 15 months long, but could vary between 60 and 615 days.

[11] On July 1, 1995 a new law concerning the total defense of Sweden came into force. It encompassed both military and civil service (*lagen om totalförsvarsplikt* 1994:1809) and stated that those clearly unable to participate in military or civil service should not be called to the testing days.

2.2. The Testing Process, the Test Office, and the Role of the Test Officiator

This paper focuses on men drafted between 1990 and 2001. At this time, each young man was called to his regional test office shortly after turning 18. The specific test date was based only on his month and year of birth, municipality of residence at age 17, and, in some cases, the expected date of high school graduation.[12] In 1990, there were six regional test offices in Sweden, each serving a specific geographic catchment area; one of these test offices closed in 1995. Test offices were required to fill troop orders placed by the military. Each office filled orders from all branches of the service, including both local military units and specific military units in other parts of the country.

The testing procedure typically took two days. On day one, groups of young men (typically from the same local area) were transported by bus or train to their regional test office. It was not uncommon for one group to come early in the day and another group to arrive somewhat later. Conscripts started the test day with an information meeting, at which they were provided procedural information about the testing process and informed of their rights and obligations. The conscripts then took part in (i) a set of written tests measuring, verbal, spatial, logical and technical ability, (ii) a telegraph test, and (iii) a series of medical and physical tests, examining their hearing, vision, strength, height, weight, blood pressure, physical condition, etc. They had an examination with a medical doctor, and (typically on day two) met with one of the test office's psychologists for an interview. The results from each test were entered into the test office's computer system and additional written information was placed in a folder carried by each draftee from station to station.

Lastly, they met with a test officiator (*mönstringsförättare*) who examined their test scores and determined whether or not a conscript would be exempted from service due to mental

[12] Carlsson et al. (2015) discuss the assignment of test dates and test offices in great detail.

and/or physical health problems. During this time period, exemptions were not discretionary. They were based on a pre-determined set of health criterion. A conscript's health scores were determined before meeting the officiator by the doctors in the recruitment office. Officiators were not involved in the process that assigned recruits their health scores.

Those who were not exempted (the vast majority) were then assigned a specific service category and preliminary start of service date by this officiator. That is, all non-exempted conscripts left the recruitment office with a specific military assignment and service commitment that they were expected to fulfill.[13]

Placement into service categories was based on an individual's full range of test scores, interviews, specific skills (e.g. driver's license, language skills) and, to some extent, the conscript's preferences for service type, year, and location. Each service category had a well-defined job description and correspondingly well-defined tasks and ranks.[14] Each service category was also associated with a set of minimum tests scores needed to be assigned to that category. For example, a specific officer position might require that a conscript received at least a 7 (on a 1-9 stanine score) on all of the major test categories. Officiators were constrained by these test score requirements when assigning service categories.[15]

Test officiators used computers to help make the first match between a candidate's scores and a set of suitable service categories. Recall that each test office had a specific troop order to fill. The computer would present the officiator with a set of potential service categories that was created by the intersection of the troop order that needed to be filled and the conscript's

---

[13] Draftees could also request to do weapon-free service. But very few requested this option. In 1994, for example, only 0.1% of all draftees requested weapon-free service (Pliktverket 1994). A slightly larger share (0.2%) stated that they were conscientious objectors and refused to do any form of military or civil service (Pliktverket 1994). These cases, however, were dealt with by the central office *after* the test day. Conscientious objectors were required to apply for an exemption in writing during the six month period after their test date.

[14] Though there is some variation over time, there were approximately 1200 service categories during most of the 1990s and 2000s (SOU 2000:21 Bilaga 3).

[15] Officiators were only allowed to deviate from these minimum test score requirements by one point on one test; and then only if the recruit had an exceptionally high score on at least one other test that was deemed particularly relevant to the position. These types of exceptions were rare at this time, since there was no shortage of qualified people to fill all of the positions.

tests scores. The officiator decided the exact service category assignment after interviewing the conscript.[16]

During the downsizing of the Swedish armed forces, nearly all service categories were oversubscribed. More people were required by law to serve than the military could place into service. So, how did the recruitment office choose who would serve and who would not serve? And how could the officiator affect the probability of service?

The most suitable persons *within* each service category were chosen to serve (SOU 2000:21 Bilaga 3 and SOU 2004:5).[17] Thus, qualified persons placed into higher ranking and more skilled categories were not always called to serve, while someone with lower qualifications and tests scores might still be called to serve since they were assigned to a different service category with different needs. Ranking of the "most suitable" candidates was not strictly based on tests scores; willingness to serve, a conscripts preferences for when and where to serve, and *the test officiators personal, subject judgment* (written down in the conscripts case file as brief notes) all played into this decision.

Importantly, since the military was continually downsizing during the time period we study, it was the officiators' responsibility to decide which individuals within each service category would be called to serve. On the margin, this left room for discretionary judgement on the part of the officiators when determining the most suitable candidates. Officiators discussed these decisions both within and between test offices on a regular basis throughout the year (typically four times a year). Thus, the test officiator can affect the probability of service of the

[16] As mentioned above, officiators were constrained by a clear set of rules from placing recruits with low test scores into high test scores positions. But they also tried not to place high test score men into low test score positions. All of the officiators we interview stressed the importance of placing the right man into the right position. We were, however, given several concrete examples of cases when the officiator did place a high test score individual into a somewhat lower test score position. In these few cases, the recruit had a clear motivation for wanting to perform a specific task and the officiator agreed that having this person in that particular position would be beneficial to the recruit and the military. Deviations from the normal matching process, however, were rare at this time (which can be seen in the data), since there was no shortage of men with the correct set of qualifications and test scores to fill each position. Nearly all service categories are oversubscribed.

[17] All of the officiators we interviewed stressed this point.

marginal draftee by (i) assigning a draftee to a service category that is in relatively high demand and (ii) advocating the case of his or her favorite candidates for the position.[18]

As such, the test officiator plays a significant role in the assignment of men to active military service during the period when not all men are required to do military service, i.e. when officiator discretion is relevant; the variation across individual test officiators in their propensity to assign men to service (both within and across test offices) is a key component of our identification strategy. Our empirical strategy also relies on the random assignment of draftees to officiators. We argue this to be the case based on the actual test day routines and anecdotal evidence from interviews with officiators working in different offices during this time period. The story is quite simple. Draftees arrived at the center, and were led through a series of test stations, some which ran in parallel and took more or less time to complete. Each conscript carried a folder with his personal information and that which was added at each test station (information is also entered into the computer system). The conscript arrived in a waiting room outside the officiators' offices (there are always multiple test officiators in each office) and placed his folder on the top of a pile in a box. The next available officiator removed the folder from the bottom of the pile and met with that conscript. This match is as good as random once we condition on test year and county fixed effects.[19]

The test officiators themselves, as well as many Swedish conscripts that we have talked to, insist upon this random match.[20] In interviews, the officiators stressed that test officers did not specialize in filling certain types of jobs or pick who to interview. All recruits had to be interviewed, and these were done on a first-come first-serve basis – the first available test officiator was matched with the next draftee in line. Furthermore, officiators did not have

[18] All test officers had been (or still were) officers in the military. Most were men. Towards the end of this time period, there were also a number of women but we do not know the specific identities of the test officiators.
[19] Recall that test office is determined by geographical location, i.e. by county.
[20] When interviewing draft officiators, we asked them how each draftee was assigned to his test officiator. The only answer we ever received was that it was as good as random (in Swedish, *slumpmässigt*), since all offices used a simple first-come first-serve que system to assign draftees to officiators.

individual quotas nor their own list of positions to fill. The test office received orders from the military for troops (number and type) and all officiators worked together to fill the office-wide order. Section 3.3 provides empirical evidence of this random matching.

2.3. Data Description

We study men born in Sweden between 1968 and 1983 who take the enlistment tests from 1990 to 2001. We have a 70% sample of these men from Statistics Sweden's Multigenerational Register (*flergenerationsregistret*), which allows us to connect these men to their parents. This data have been matched to data from The Swedish Military Archives (*Krigsarkivet*), The Swedish Military Recruitment Office (*Rekryteringsmyndigheten*), The Official Convictions Register (*belastningsregitsret*), and various register data from Statistics Sweden using each individual's unique personal identification number.

2.3.1. The Draft Data

We have draft board data from The Swedish Military Archives and The Swedish Military Recruitment Office, though we only use the latter in our empirical analyses. The former was used to help characterize historical trends in testing and service (see Figure 1).

Our main analysis uses the test officiator id as an instrument for military service for potential conscripts who tested between 1990 and 1996 for two reasons. First, the officiator IV strategy does not work for earlier cohorts, since almost everyone who tested before 1990 served; that is, there is little room for officiator discretion. Second, the officiator id is missing from the recruitment office's records for cohorts tested between 1997 and 2001.[21]

[21] The officiator variable re-appears in the data in 2002. But by this time, less than half of all Sweden born males were placed into service categories and less than half of those were enlisted in the military (see Figure 1). Although it was still illegal to refuse military service, it had become more or less optional for young men.

The draft board data cannot be used to assign treatment status, since the data are incomplete when it comes to identifying who actually served in the military. Thus, to identify treatment status, i.e. those who served at least two months, we turn to the national tax registers. Every conscript who served for at least two months received a small taxable income from the government, which is specially marked in the tax register on an annual basis. Since we see that a payment was received during the year but not when it was received, we can only use this tax data to identify individuals who were enlisted but not the exact dates of enlistment. We do, however, have exact service dates for those who tested from 1997 to 2001, which we use to perform an alternative (non IV) incapacitation analysis.

Finally, a number of additional variables are used from the draft board data, primarily for descriptive purposes, tests of identification, or robustness checks. For the 1990 to 1996 cohorts, we use the test date, test office, height, weight, bmi, general ability test scores (stanine scores, 1-9), physical capacity (stanine scores, 1-9), health categories, psychological capacity (stanine scores, 1-9), and whether the person was assigned a service category. Health and physical capacity scores are both summary measures based on a series of underlying tests.

2.3.2. Outcome Variables – Crime

The first order aim of this paper is to study the causal effect of mandatory military service on crime. To this end, our data were matched with the official crime register (*belastningsregistret*) for Sweden by the National Council for Crime Prevention (BRÅ), providing a full record of criminal convictions from 1973 to 2012. As is typical with administrative crime data, we cannot directly observe criminal behavior, and rather, use convictions as a proxy for criminality. For each conviction, we observe the type of crime and the date the offense was committed. We study overall crime, *Any Crime*, and by six specific crime categories: *Weapons*, *Violent*, *Traffic*, *Theft*, *Other*, and *Drugs & Alcohol*. We define extensive margin crime variables that are equal

to one if the individual has at least one conviction in the appropriate category. At the intensive margin, we look at both the number of convictions as well as dichotomous variables indicating whether individuals have two or more and five or more convictions. We use the latter set of variables to investigate and alleviate concerns that the intensive margin results are driven by a few individuals with an extremely large number of convictions. Since we know when each crime was committed, we create crime categories based on age and classified as pre-service (ages 15-17), during service, and post-service (ages 23-30). The non-crime outcome variables are described when presenting the results.[22]

2.3.3. *Background and Control Variables*

We also use a number of background and control variables from register data held by Statistic's Sweden. We make use of *Birth Month*, *Birth Year*, and *County* of residence at age 17.[23] We record if a person was enrolled in a 2- or 3-year high school program, since this was used in some cases to help assign test dates, and create measures of mother's and father's education and income to ascertain the socioeconomic background of our draftees.[24]

3. Identifying a Post-Service Effect

The empirical analysis is conducted in two stages. The first utilizes the 1990 to 1996 test cohorts, for whom we can identify the officiators who reviewed the draft board test results and made service decisions. We identify the causal effect of service on post-service crime using an instrumental variable design that capitalizes on the random assignment of potential conscripts

[22] Military courts were abolished in Sweden in 1949. So all criminal infractions, regardless of whether they are committed on or off base, are handled by the local police and criminal justice system.
[23] We also have parish and municipality at age 17.
[24] Education is measured in seven levels. Income is measured as the log of average income using all available income data from 1968-2012. The income concept used here is pre-tax total factor income.

to officiators who assign more or less individuals to service.[25] Though we can use this design to identify post-service crime effects that are not muddied by incapacitation, we cannot directly identify the incapacitation effect itself for this sample given the unavailability of the exact dates of service. Thus, Section 5 uses the 1997 to 2001 test cohorts, for whom we know the exact dates of service, in a matching framework to further isolate incapacitation. This section describes the use of the officiator as an instrument.

3.1. Officiator Assignment as Instrument for Military Service

The primary aim of this section is to identify the causal effect of conscription on post-service crime and labor market outcomes. To that end, consider a regression that relates an outcome of interest, y_i, for individual i to whether he was conscripted into the military, $Conscript_i$.

$$y_i = \propto + \beta Conscript_i + X_i + \varepsilon_i \tag{1}$$

Even with a large set of observable controls, X, conscription is likely to be correlated with the error term due to the selection process. Because the tests themselves, as well as unobservable determinants of the results (like background, ability, performance under pressure, etc.), affect the likelihood of service as well as crime and labor market performance, Ordinary Least Squares (OLS) estimation of equation (1) will yield biased estimates of the effect of conscription.

Thus, we propose to instrument for *Conscript* with a dummy variable indicating whether the individual is assigned to a '*high service rate officiator*'. This variable is equal to one if the annual share of testees assigned to the officiator who serve is greater than the national share who serve in that test year.[26] That is, is individual i assigned to an officiator with a relatively high service rate in test year t, compared to the average service rate in test year t? As such, we

[25] In spirit, the design is similar to using randomly assigned judges (Kling, 2006; Aizer and Doyle, 2015; Mueller-Smith, 2015) or investigators (Doyle, 2008) as exogenous sources of variation for sentences and foster care, respectively.

[26] We create this variable based on our baseline sample of all non-immigrant males who tested each year, and not the final analysis sample.

are utilizing the exogenous variation in the chance of service given the officiator one is assigned in a given year from the pool of potential officiators in that year, rather than variation that arises due to some officiators working at the beginning of the sample period (with somewhat higher service rates) versus the end.

Though there are clearly alternative ways to define the instrument, such as officiator fixed effects (i.e. the leave out mean), we use a dichotomous variable because of its simplicity and ease of interpretation (especially in the first stage). In contrast to other papers using harsh judges as instruments for sentence severity (Aizer and Doyle, 2015; Kling, 2006), the first stage is strong enough to allow for such a simple specification; however, we also demonstrate the robustness of our results to an alternative instrument – namely the leave out annual mean. Note again that it is important to use the leave out *annual* mean rather than simply the leave out mean for the officiator because of the downward trend in the number of service positions and not all officiators are observed in the same years.

To isolate the post-conscription effect of service from incapacitation (and be certain that outcomes are observed after service is completed), we (i) define our primary outcome variables by age, emphasizing crime between ages 23 and 30 and labor market outcomes between ages 23 and 34, and (ii) restrict the sample to individuals who completed service before age 23. Before turning (in Section 3.3) to the relevance and validity of our proposed instrument, we briefly discuss how our analysis sample is created and descriptive statistics.

3.2. Sample Creation and Descriptive Statistics

Our baseline data set consists of 231,583 non-immigrant males born from 1964 to 1990, who tested from 1990 to 1996, i.e. when the officiator variable is available. We omit about 15,000 individuals assigned officiators with less than 100 cases in their test year and slightly more than 40,000 individuals who were missing health group information (about 1500 individuals) or

assigned to health groups that 'never' serve in a given year. It is important to note that the decision to assign individuals health categories that are ineligible for service is not made by the officiator, nor is the decision that these health categories will not serve in a given year. Thus, these individuals are omitted from the baseline sample since it is impossible for the officiator to influence whether or not they serve; we will return to this sample that is ineligible for service for health reasons in a falsification exercise later in the paper. After dropping less than 1,000 individuals who are not assigned to service (i.e. they are listed as ineligible to serve, though we do not know the reason) and less than 5,000 individuals who are 23 or older in the year they finish service (or for whom the year is unknown), we obtain our final sample of 1686,818 non-immigrant males tested between 1990 and 1996.[27] See Appendix Table 1 for details.

Our data contain 67 officiators in the six primary test offices in Sweden. In any given year, the number of officiators observed is between 25 and 29 (except 1993 when we observe 37 officiators). The average number of officiators in each test office and year is about 10, since some officiators are not stationed in a single office but rotate across test offices in a given year. In fact, just 42 percent of officiators are stationed in a single office each year; 19 percent in two, 17 percent in three, and the remaining 22 percent in four or more.

Table 1 provides summary statistics for the analysis sample – overall, by service, and by assignment to a high service rate officiator. During this period, 75 percent of the sample that was eligible to serve actually served. Looking across test years in our analysis sample, Figure 2 shows that service rates begin to decline after the 1993 test year. (As we already condition our sample on eligibility to serve, service rates for the entire sample would be even lower.)

Table 1 also demonstrates that despite coming from comparable birth and test cohorts, those who serve are 21 percentage points more likely to have been assigned a high service rate

[27] Specifically, for these cohorts, we know the last year in which they are observed receiving income from the military according to tax records. We drop individuals who are 23 or older in this last year.

officiator than those who do not serve. This pattern is clearly suggestive of the first stage relationship that we intend to use in the instrumental variable analysis.

The middle panels of Table 1 characterize the potential conscript's offense specific criminal history and socioeconomic status prior to testing, as well as his performance on the test day. The service sample is positively selected in all dimensions: they have less criminal history (for all crime categories), come from more educated families, are more likely to attend three versus two year high schools, and have higher ability, physical capacity, and especially psychological capacity test scores. Though those who serve are less likely to have a criminal history, having a criminal history does not disqualify one from service. In fact, 13 percent of the service sample has at least one conviction (in any crime category) prior to age 18. This positive selection into service is observed regardless of who assigns service – i.e. high or low service rate officiators.[28]

Finally, the bottom of Table 1 considers the main crime outcomes. Overall, we see that 10 percent of the sample is convicted of at least one crime between ages 23 and 30. And though the average number of crimes convicted (including zeroes) is 0.31, just four percent of the sample has two or more convictions between ages 23 and 30 and just one percent has five or more convictions. At both the extensive and intensive margins (not shown), the largest crime category is traffic offenses; six percent of the sample has at least one traffic conviction. The other crime categories have much lower conviction rates: one percent for weapons, two percent for violent, one percent for theft, two percent for drugs, and two percent for other offenses. We also see that for almost every crime measure, the crime rate is lower for the service sample than the non-service sample. The above described positive selection into service, however, makes it clear that this cannot be interpreted as anything more than a correlation.

[28] Available from the authors upon request.

3.3. Instrument Relevance and Validity

A good instrument is one that is relevant and valid, and for which the assumption of monotonicity holds. In the current context, this implies that (i) assignment to an officiator with a higher than average annual service rate significantly increases the likelihood of serving, (ii) whether a testee is assigned a high service officiator is unrelated to testee unobservable characteristics, and (iii) officiators do not dramatically change their behavior and switch from being a high service officiator to a low service officiator (or vice versa) depending on the type of testee they meet; that is, all testees must be at least weakly more likely to serve when facing a high service rate officiator.[29] We argue that the proposed instrument meets all three criteria.

With regards to validity, recall from Section 2 that testees are randomly assigned an officiator after completing their battery of tests.; this implies that both observable and unobservable testee characteristics should be uncorrelated with officiator characteristics, including whether they are a high service rate officiator. However, since officiator characteristics may vary *across* test offices or years in a way that is correlated with testee characteristics, we argue that conditional on when and where the individual was assigned to take the test, the officiator to which a testee is assigned is as good as random. We use county fixed effects to control for the fact that test office assignment is based on where one lives.[30] We believe that county fixed effects to be more appropriate than test office fixed effects themselves since (i) individuals from the same region (and birth month groups) were often bussed to the test offices together and (ii) it was common for the same officiators to be in multiple test offices. However, we demonstrate that our results are completely robust to controlling for test office.

As a first crude test of random assignment, the last two columns of Table 1 present the raw sample means for individuals assigned to high and low service rate officiators, respectively.

---

[29] See Mueller-Smith (2015) for an in depth discussion of the monotonicity assumption in the context of the judge fixed effects identification strategy.

[30] While parish of residence is officially what is used in assigning test offices, this unit is too small to conduct this analysis. However, we provide evidence that county is sufficient to achieve conditional random assignment.

While the percent difference in the share serving is 16 percentage points (almost 19 percent) across these two groups, the raw percent differences in the observable characteristics are in many cases less than five percent.[31] It is not surprising, however, that some observable differences (albeit small) do exist, as the raw statistics do not condition on time and geography.

Table 2 more formally tests for random assignment by regressing assignment to a high service rate officiator on the full set of test day (height, weight, bmi, and ability, physical capacity, and psychological test scores) and pre-test day (offense specific crime before 18, mother and father schooling and income, and 2-year versus 3-year high school) characteristics. All tests of significance are based on standard errors clustered at the officiator level. To summarize the results, we tabulate the number of significant coefficients and present p-values for an F-test of the joint significance of the pre-test day and test day characteristics. When excluding all fixed effects in column (1), these variables are jointly significantly different than zero; this is driven especially by the pre-test day characteristics. Even in this raw data, however, there is strong evidence of random assignment. (i) There is little difference in test day characteristics for high and low service officiators. (ii) It is not always those with better or worse backgrounds who are consistently more likely to be assigned a high service officiator. (iii) The estimated magnitudes are relatively small, given that 50 percent are assigned a high service officiator. (iv) All controls only explain two percent of the variation in assignment.

Nevertheless, we proceed by demonstrating that the observed relationships get an order of magnitude smaller and become jointly insignificant when controlling for county and test year fixed effects. Specifically, adding county fixed effects in column (2) decreases the number of significant coefficients (including dummies for missing observations) from 12 to six and results in a p-value of the joint significance of the pre-test and test day controls of 0.08.[32] Further

[31] One exception is the share in a 2-year high school, but as 2-year programs are greatly trending down in this period (and 3-year programs trending up), this raw comparison may not account for this variation.
[32] Controlling for county fixed effects eliminates the significance on almost all dummies for missing data, suggesting that the missing observations in the controls are occurring selectively by geography.

controlling for test year fixed effects in column (3) reduces the number of significant coefficients to three and yields a p-value on the joint test of 0.25. Adding county by test year fixed effects in column (4) further reduces the magnitude of the coefficients (just one is significant at the five percent level), and results in a p-value of 0.14. Finally, column (5) of Table 2 examines how that variation in the high service officiator assignment that is *unexplained* by county by test year fixed effects is related to the pre- and test-day characteristics; that is, the dependent variable is the residual from a regression of a high service officiator dummy on county by test year fixed effects. None of this unexplained variation is explained by the full set of observable controls.

We thus argue that officiator assignment is random when conditioning on county by test year fixed effects, and define our baseline instrumental variable specification accordingly (i.e. include test year, county, and test year by county dummies). Of course, we cannot rule out non-random assignment based on unobservable characteristics. But, together with the anecdotal evidence of random assignment from officiator interviews, we believe this makes a strong case for the validity of the instrument. We provide further evidence of random assignment throughout the paper – namely that the baseline first stage and instrumental variable estimates are not sensitive to a large set of controls. We also use the sample of individuals ineligible for service because of health reasons to conduct a falsification exercise; for this sample, we examine the reduced form specification of whether assignment to a high service officiator affects crime outcomes. It could only do so if random assignment fails – i.e. testees are assigned officiators according to characteristics related to future crime.

Table 3 presents the results of the first stage analysis: does assignment to a high service rate officiator increase the likelihood of serving in the military? According to column (1) with no controls, assignment to a high service rate officiator significantly increases the likelihood of service by almost 16 percentage points. However, much of this is accounted for by regional and

temporal variation in service rates. Column (3) includes test year by county fixed effects – in this case, assignment to a high service rate officiator increases the likelihood of service by 7.9 percentage points, with an associated F-statistic of 32 (well above the weak instrument threshold). Column (4) demonstrates the stability of this first stage relationship to controls for (i) test office and test office by test year fixed effects, (iii) pre-test day characteristics (including the full set of crime specific history variables), and (iv) test day characteristics (including ability, physical capacity and psychological test scores). This has little impact on the magnitude of the first stage effect and associated F-statistic.

Columns (5)-(8) of Table 3 present results in support of the monotonicity assumption. Specifically, we re-estimate the first stage relationship by subsamples: criminal versus no criminal history and low versus high education fathers. Regardless of subsample, assignment to a high service rate officiator significantly increases the chance of service seven or more percentage points. Those with and without a criminal history are 9.7 and 7.1 percentage points more likely to serve, respectively. Similarly, those with low and high education fathers are 8.9 and 6.9 percentage points more likely to serve. Note that the estimates for the high and low background subsamples do not significantly differ from each other.[33]

Before turning to the baseline results, we further investigate whether assignment to a high service officiator affects more than whether an individual simply serves or not. Does it also affect the type of service? Our ability to interpret the IV estimates as the causal effect of military service relies on the assumption that it is only service that is affected. The results in Table 4 clearly support this assumption. Assignment to a high service officiator does not significantly affect the branch of service (army, navy or air force), the rank of service (private,

[33] To further investigate the assumption of monotonicity, we re-calculated our high service officiator dummy for each year and officiator using samples of (i) only men from low SES families, (ii) only men from high SES families, (iii) only men with prior convictions, and (iv) only men without prior convictions. We then examine whether any officiator (in a particular year) changes from a high to low service status (or vice versa) when facing one of these four types of men. They do not. Only a few year by officiator dummies are re-categorized and no officiator is consistently re-categorized in multiple years when facing a particular type of young man.

corporal, or sergeant/2nd lieutenant) or whether in a combatant position. Note that rank can also be a proxy for length of service, as higher ranks serve more months. Compared to an F-statistic of 32 on whether an individual serves at all, the F-statistics characterizing the type of service are between zero and three. Though we still present the reduced form baseline results in the next section, we believe these results justify our emphasis on the instrumental variable analysis.

4. High Service Rate Officiator Instrumental Variable Results

4.1. Baseline Results and Robustness Checks

Table 5 looks at the relationship between serving in the military and overall post service crime from ages 23 to 30; Panel A considers the extensive margin (at least one conviction) while Panels B – D consider the intensive margin (number of convictions, two or more convictions, five or more convictions). OLS estimates in Columns (1) – (3) find a negative correlation between service and overall crime, regardless of how crime is measured. When including county by test year fixed effects in column (2), service is associated with a 1.7 percentage point (17%) reduction in the likelihood of at least one conviction, 0.13 (42%) fewer convictions on average, and 1.3 and 0.6 percentage points (31% and 46%) lower chance of being convicted of two or more and five or more crimes, respectively. Though controlling for test office fixed effects and the full set of pre-test and test day characteristics in column (3) substantially reduces the magnitudes of these relationships, they remain negative and highly significant.

Columns (4) and (5) of Table 5 present the results of the reduced form regressions of each of the four post-service crime outcomes on assignment to a high service officiator. Our baseline specification, with just county by test year fixed effects, is presented in column (4). Assignment to a high service rate officiator significantly increases the chance of at least one conviction between ages 23 and 30 by about 0.7 percentage points (7%) and the average number of convictions by 0.04 (14%). Controlling for the full set of observable controls in column (5)

has little impact on these point estimates, but increases the precision with which they are measured. Though effects of similar magnitudes are seen for convictions of two or more and five or more crimes (13% and 12% respectively), these estimates are not significant.

Columns (6) and (7) of Table 5 present the 2SLS estimates of the relationship between military service and post-service crime, using assignment to a high service officiator as an instrumental for military service. In other words, we scale the reduced form estimates in columns (4) and (5) by the first stage estimates in Table 3 (approximately 0.079 without controls and 0.075 with controls) to identify the causal effect of service on crime. All instrumental variable specifications cluster the standard errors on the officiator. These results indicate that military service in fact has a large positive causal impact on post-service crime from age 23 to 30. Serving in the military increases the chance of post-service conviction by 8.9 percentage points and, on average, leads to an increase of 0.56 convictions. Relative to the mean, these effects are quite large – an increase of about 90 percent at the extensive margin and more than 100 percent at the intensive margin. Using the dichotomous intensive margin variables, we see comparatively large, though insignificant, point estimates. Though the associated standard errors and confidence intervals are quite large, the IV estimates clearly indicate a significant positive effect of service on crime. One should also keep in mind that this is the local average treatment effect (LATE) or the effect of service on young men to whom officiator assignment matters; it does not seem infeasible that service differentially affects these young men.

Our baseline specification clusters on officiator. However, one could argue that it is more appropriate to cluster in other dimensions, such as test year or test year and officiator. Appendix Table 2 demonstrates the robustness of our results to clustering the standard errors on county (column 2), both officiator and test year (column 3), and officiator by test year (column 4). Columns (5) and (6) of Appendix Table 2 presents the results of using the leave out annual mean as the instrument, i.e. the actual share of testees who serve in a given test year

(excluding individual *i*). The point estimates are quite comparable and somewhat more precisely estimated: military service leads to a 5.7 percentage point increase in the likelihood of conviction and, on average, an increase of 0.55 convictions. Column (6) demonstrates the robustness of the leave out annual mean instrument results to including the full set of controls.

These results clearly demonstrate a positive impact of military service on *overall* crime. But, one may ask whether such an effect is seen across crime categories or driven by a particular category. Table 6 addresses this question. The dependent variable in Panel A is whether one has at least one conviction in the crime category listed at the top of the column: weapons, violent, traffic, theft, drugs and alcohol, and other offenses. The dependent variables in Panels B and C measure the intensive margin: the number of convictions and whether there are two or more convictions in each crime category. We do not consider whether there are five or more convictions by crime type given the extremely low sample shares with five or more convictions of the same crime. Row (a) of each panel presents the baseline specification; row (b) adds the full set of controls, including have a criminal history in each crime category.

At the extensive margin, service significantly increases the likelihood of at least one weapons conviction between ages 23 and 30 by 1.7 percentage points and a violent offense conviction by 4.9 percentage points. At the intensive margin, service significantly increases the number of convictions for: violent offenses by 0.10, theft by 0.17, drugs and alcohol offenses by 0.12 and other offenses by 0.10. Service also significantly increases the chance of two or more convictions for theft and drugs and alcohol offenses. When considering the two intensive margin results together in Panels B and C, one can conclude that it is only theft and drugs for which we truly see an intensive margin effect, which is not completely driven by outliers in the number of convictions. The intensive margin effect on violent crimes would appear to be largely driven by the extensive margin, i.e. the first violent crime conviction. As we saw when looking

at overall crime, these effects are quite large – often more than two times the mean. Though the remaining estimates are in significant, they are all positive and also quite large.

Appendix Table 3 replicates row (a) of Table 6 using the leave out annual mean instrument and yields qualitatively similar results; the primary difference is the additional finding of a significant increase in traffic offenses at both margins. Finally, controlling for the full set of controls in row (b) of Table 6 has minimal impact on the magnitude or precision of the effects. Given that some variables, especially criminal history, are particularly strong predictors of future crime, the insensitivity of the IV estimates to their inclusion additionally supports the identifying assumption that testees are randomly assigned to officiators.

4.2. Heterogeneity Analyses

This section examines whether the results are heterogeneous across two dimensions of the testee's pre-test background: (i) having a criminal history prior to age 18 (Panel A of Table 7) and (ii) coming from a low socioeconomic status family, for which we use father's education as a proxy (Panel B of Table 7). Appendix Table 4 presents the same specifications using the leave out annual mean instrument. For the most part, we find that the results are driven by individuals from disadvantaged backgrounds – with at least one pre-service conviction or with low education fathers (nine or less years of schooling). As seen in Table 7, serving in the military increases the likelihood of any convictions by 22 percentage points (or slightly less than 100%) for those with a criminal history. Large point estimates are seen for all crime categories, thought they are only significant for violent offenses and thefts. For the low education father sample, military service increases the chance of conviction by 11 percentage points overall, with significant effects seen for weapons, violent and drugs and alcohol offenses.[34] For those with no criminal history or high education fathers (more than nine years

[34] When using the leave out annual mean instrument, similarly large and positive coefficients are seen for those with a criminal history, but the overall effect is not significant; marginally significant coefficients are seen on

schooling), the offense-specific point estimates are often an order of magnitude smaller and generally insignificant.[35]

4.3. Falsification Test: Additional Test of Random Assignment

This section presents a falsification test, which can be interpreted as an alternative test of random assignment, using both the overall sample and heterogeneity sub-samples. Panel A of Table 8 presents the reduced form effect of being assigned a high service officiator on having any convictions between ages 23 and 30 for our main analysis sample; columns (1) – (5) include the baseline controls while columns (6) – (10) include the full set of controls. Consistent with the 2SLS heterogeneity results, assignment to a high service officiator for the analysis sample significantly increases the chance of a post-service conviction (at the 5 or 10% level) overall and for those with disadvantaged backgrounds in terms of criminal history and father schooling. In Panel B of Table 8, we present the same specifications for a sub-sample of individuals who were assigned officiators but did not serve for health reasons.[36] For these individuals, assignment to a high service officiator could only affect crime from ages 23 to 30 if: (i) the short meeting with the officiator actually had a direct effect on the testee (which seems rather unlikely) or (ii) the testees were not randomly assigned to the officiator to start with. In the latter case, if individuals were non-randomly assigned officiators on the basis of characteristics that are correlated with future crime, then one could see a significant reduced form relationship. However, as seen in Panel B of Table 8, the assignment of a high service officiator is not

weapons offenses and thefts. For the low father education sample, the same offense specific coefficients are significant with the leave out annual mean as the high service officiator instruments.

[35] Two exceptions are seen: violent offenses for the no criminal history sample when using the high service officiator instrument and traffic offenses for the high father education sample when using the leave out annual mean instrument. These results are not robust to the choice of instrument however.

[36] Note that this analysis is conducted in a reduced form rather than IV framework since our falsification sample by definition does not serve.

associated with future crime for this sample of individuals who are ineligible for service. This provides additional support of the random assignment of officiators.

4.4. Non-Crime Outcomes

This section considers the effect of service on non-crime outcomes. Not only does service occur near the peak of the age-crime profile, but it is also at a time when young men are embarking on either higher education or entering the labor market. To this end, we create a set of non-crime outcome variables from Statistics Sweden registers and the tax registers. From the latter, we have information on income and the use of benefits, including sickness, disability, welfare, and unemployment. *Income* is the log of average income between the ages 30 and 34.[37] The income concept we use is pre-tax total factor income. *Sick Days* is the number of days a person has received sickness benefits between the ages of 23 and 34. *Disability Pension* is the number of years in which a person has received a partial or full disability pension between the ages of 23 and 34.[38] *Unemployment Benefits* equals the number of years during which an individual has received at least one payment from the unemployment insurance system between ages 23 and 34. *Months on Welfare* is the number of months a person has received welfare benefits between ages 23 and 34. We also consider the extensive margins for unemployment, welfare, and disability pensions. Lastly, we create a dichotomous variable for *Education*, which equals one if an individual has obtained at least some college by 2012 and zero otherwise.[39]

---

[37] Measuring income for Swedish men at these ages has been shown to be a reasonably good proxy of their permanent income (Böhlmark and Lindquist 2006). We would have preferred to use a measure of income averaged over a longer series of incomes measured and at later ages, e.g. 30 – 40. Unfortunately, our cohorts are simply too young. So we do not have the data to do this. Using income measured before age 30 would severely bias our measure of income, since high skilled workers have not yet reached their earnings potential. It may even appear as if their income potential is lower than that of low skilled workers.

[38] Full (partial) disability pensions are granted to those who have no (reduced) work capacity due to mental or physical health issues that are deemed permanent. That is, the disability pension system tends to be an absorbing state. Those with health issues deemed temporary receive sickness benefits when they are unable to work.

[39] We focus on legitimate labor market outcomes, as these are a natural complement to crime, or the illegitimate labor market. We considered family outcomes, but do not find any significant or consistent effects of service on partnerships or marital status. We have not studied fertility, since we do not observe completed fertility.

Table 9 presents the baseline instrumental variable results for the whole sample, the criminal history versus no history samples, and the low versus high education father samples. Overall, and for each subsample, we do not find any evidence of a causal effect of service on attaining more than 12 years of schooling. However, we do find evidence that individuals from disadvantaged backgrounds (have a criminal history or low educated father) are made worse off in the labor market as a result of military service, while those from better backgrounds are either not affected or made better off. Specifically, with regards to income, military service significantly decreases log income between the ages of 30 and 34 by 2.3 percent for individuals with low educated fathers and increases incomes by 1.7 percent for those with high educated fathers; similar estimates (though not quite significant) are seen when splitting the sample by criminal history.[40] We also find that military service increases the likelihood of receiving any unemployment benefits from 23 to 34 by 36 percent for those with a criminal history and 28 percent for those with low educated fathers; there is no effect for those with no history or higher educated fathers. The same pattern is seen when looking at the number of years of unemployment, though the estimates are generally insignificant. Military service also significantly increases the number of months of welfare receipt from ages 23 to 34 by more than eight months for those with a criminal history and almost five months for those with low educated fathers; these effects are quite large relative to the respective sample means of 4.2 and 1.98 months.

Appendix Table 5 replicates the non-crime outcome analysis using the leave out annual mean as the instrument. While some of the point estimates differ, the same qualitative story is seen: military service leads to worse labor market outcomes for those from disadvantaged backgrounds and better outcomes for those at the upper end of the distribution. Using the leave

[40] The standard deviation of income is 0.78. Thus, the difference between these two estimated effects (0.21+|-0.29| = 0.5) is equal to 64% of a standard deviation.

out annual mean, we even see that service decreases the chance of higher education for those with a criminal history.

We see less evidence of a heterogeneous impact of military service when looking at disability benefits and the number of sick days; if anything (especially using the high service officiator instrument), we see some evidence that military service actually leads to an improvement in outcomes for everyone.

Taken together, these results suggest that peacetime conscription increases participation in the illegitimate labor market and decreases participation in the legitimate labor market, particularly for the most disadvantaged individuals. This contrasts the belief/hope that providing discipline to individuals already at risk for a life of crime will put them on a better path and that the human capital skills gained during conscription improve the labor market outcomes for those coming from a disadvantaged starting point. Unfortunately, it is not possible to empirically disentangle the channels through which these effects occur. Military service clearly impacts both post-service crime and labor market outcomes; but it is unclear whether service has a direct effect on one of these outcomes, which indirectly affects the other.

On the other hand, there is some evidence that all individuals are healthier as a result of peacetime conscription, perhaps because of the physical training received. Of course, one should recall that this sample is already positively selected on health status. This, not surprisingly perhaps, contrasts previous research that finds detrimental effects of Vietnam Era service on a number of morbity measures (Johnston, Shield, Siminski, forthcoming) and disability receipt (Autor, Duggan, and Loyle, 2011; Angrist, Chen, Frandsen, 2010).[41]

4.5. Isolating Incapacitation

[41] Other papers that look at the impact of military service on health outcomes include Dobkin and Shabani (2009) and Bedard and Deschenes (2006).

This section uses the same instrumental variable approach to identify incapacitation, i.e. whether individuals commit less crime *during* military service. Because we do not know the exact dates of service for our IV sample, we (i) restrict the analysis to individuals tested in the year they turned 18, almost all of whom serve when 19 and/or 20, and (ii) create measures of crime convictions at 19 and 20. But, we cannot rule out that some of these crimes occurred before or after service, given that service is not for the whole two year period.

Panel A of Table 10 presents the extensive margin results for the whole sample, and by criminal history. For the whole sample, service decreases the likelihood of any at age 19 and 20, though this is insignificant. Looking across crime categories, about half of the estimates are negative (in contrast to all positive coefficients in the post-service analysis); there is even a marginally significant negative effect of service on violent crime convictions. This 'incapacitating' violent crime effect is seen for those with and without a criminal history. Individuals with a criminal history are also significantly less likely to be convicted of traffic offenses while individuals with no history are significantly less likely to be convicted of drugs and alcohol offenses (though the conviction rate is extremely low for this sample) and more likely to be convicted of traffic offenses; the latter is the only significant positive estimate.[42]

The results generally suggest that military service has some incapacitating effect on crime. Panel B of Table 10 supports this by looking at the effects of service on crimes from ages 23 to 30 for this selected subsample of individuals tested at age 18. We still see positive coefficients across the board, and a large, significant positive coefficient for violent offenses; thus, the negative violent crime effect observed at ages 19 and 20 cannot be attributed to the sample of 18-year old testers being non-representative of the whole sample.

[42] This could potentially be capturing post-service effects on traffic offenses, or even traffic offenses committed during service, if individuals drive to and from service (on a weekly basis in some instances).

5. An Additional Incapacitation Analysis Using Exact Service Dates and Exact Crime Dates in a Matching Framework

Section 4.4 finds evidence suggestive of incapacitation in our instrumental variable framework. As we do not have exact dates of service for this sample, it cannot be ruled out that these crime outcomes include some pre- or post-service crimes. In this section, we take advantage of the fact that we have exact dates of service for the cohorts tested between 1997 and 2001 and apply a matching strategy to cleanly estimate the potential incapacitation effect of military service.[43] Specifically, treated individuals (i.e. those who serve) are each matched to one specific control individual (who does not serve). Each control individual is then assigned the exact same service dates as his treated counterpart. This enables us to construct the counterfactual time of incapacitation for the control group. We then use this sample in a matching framework to identify the incapacitation effect of military service.

The matched sample is constructed as follows. We start with all men who tested between 1997 and 2001. Then we keep only those men who served in the military (48,453) and those who did not serve but were assigned to a service category (72,763). Of those who serve, we keep the 28,551 (59%) for whom we have exact service dates.

During the conscription, each person is assigned a health and physical aptitude category. At this time, only those with categories A, B, D, E, F, and J actually served. We, therefore, exclude those who were assigned some other health category. This reduces the size of the potential control group from 72,763 individuals to 36,401 individuals. We drop individuals from the treatment group with obvious mistakes in their service dates and those who serve for less than two months or more than 24 months (the latter have chosen to stay on as professional military officers). This leaves us with 28,263 treated individuals who served in the military (and 36,401 potential controls).

[43] Recall that we do not have the test officiator identification number for these years. So we cannot use our preferred IV estimation strategy along with these exact dates.

We then estimate a propensity score for military service using a logit model that includes: mother's and father's education and income, enrollment in a 2-year or 3-year high school program, verbal and general ability tests scores, bmi, physical capacity and health group. We also include detailed, pre-service crime variables. We include a full set of extensive margin crime variables by crime type and age (15 to 18).

We match exactly on birth year, to insure that the treated and control individuals in each matched pair are the same age. We then use the estimated propensity score to conduct a 1 to 1 nearest neighbor matching (within each birth year cell) without replacement using a very fine caliper. This exercise produces a sample of 3,683 matched pairs of treated and controls (and a total sample of 7,366 individuals). The small sample size is the product of our strict matching routine that requires that *all* pre-service crime variables, family background characteristics, and test-day scores balance across treatment and control groups.

The young men in our matched sample are much more positively selected than those in our IV sample. They are more likely to be enrolled in a 3-year high school program, have more educated parents, are much less likely to have a teenage conviction, and have higher ability test scores than our baseline IV sample (see Appendix Table 6). We need to keep these important differences in mind when comparing results across these two different samples.[44]

In Figure 3, we plot the age crime profiles of the treated and control groups for aggregate crime at the extensive margin. Similar plots are shown for our six crime categories in Figures 4. Note that for the moment we are not making use of exact service dates; we are simply looking at age-crime profiles. The typical age during service is between 19 and 22. These ages are depicted by the two vertical lines in each graph.

[44] In an earlier version of this paper (Hjalmarsson and Lindquist 2016), we reported the results from using a difference-in-difference strategy on a much larger sample that was *not* matched on pre-service crime. Despite using this very different approach on a much larger sample, we obtained results quite similar to those reported here. That is, we always conclude that there are significant incapacitation effects.

In Figure 3, we see that the treated and control groups have very similar age-specific conviction rates up to (and including) age 18. Then, at age 19, the conviction rate for the control group goes up. This is because crimes related to drugs, alcohol and traffic are trending upwards (see Figure 4); Among other things, this trend is due to the fact that young adults are allowed to buy alcohol in restaurants and bars starting at age 18 and that they can obtain a driver's license after turning 18. The same upward jump in convictions is not seen for those who join the military, since (we argue) they are subject to a strong incapacitation effect. Although Figure 3 is not based on exact service dates, it does support the idea of an incapacitation effect, since convictions of these two groups differ quite a bit between the ages of 19 and 22, when most men typically perform their military service. In Figure 4, we see indications of this incapacitation effect across all six crime categories.

Estimates of the incapacitation effect are shown in Panel A of Table 11.[45] The estimated incapacitation effect for overall crime is negative, statistically significant and quite large (a reduction of 50% when compared to the mean). This overall effect is driven mainly by a reduction in theft and in drug and alcohol related crimes. For men with no prior criminal history, we also see a significant reduction in serious traffic crimes. Importantly, incapacitation effects can be seen for all men, both with and without a prior criminal history. Taken together, our matching and IV estimates imply that there are indeed significant incapacitation effects associated with military service.

Panel B in Table 11, demonstrates the fact that this reduction in crime is only temporary (while in service). It is not the case that the treatment group is made up of relatively low crime type men who always commit less crime. In fact, there are few significant differences in criminal behavior from age 23 to 30. The treated do commit more traffic crimes, but they also

[45] We only report the Average Treatment Effect on the Treated (ATT). None of the Average Treatment Effects (ATEs) differ by any meaningful amount.

commit fewer drug and alcohol related crimes (although this effect is only ¼ the size of the incapacitation effect).

So why are not all of these post-service effects large and positive? In Section 4, we demonstrated that it was the marginal conscript with prior criminal histories or who came from a lower SES background that responded negatively to military service. In Panel B of Table 11, we report the average treatment effect on the treated for a group of strongly positively selected individuals; hence, in line with our findings in Section 4, we should not expect large increases on post-service crime for these men.[46]

6. Discussion of Potential Mechanisms Including the Role of Peers

Military service affects the lives of young men in many different ways. In our paper, we have shown that it significantly lowers the number of sick days and disability claims among all men who serve regardless of social background. This may be due to the additional physical activity and training young men receive in the military. We also saw significant and positive income effects for young men from advantaged social backgrounds. This finding echoes that of Grönqvist and Lindqvist (2016), who show that officer training can raise the probability of becoming a manager later in life and improve wages.[47] They argue that officer training improves leadership-specific human capital. They also find positive effects on educational attainment among those who take officer training and argue that this may be due to peer effects. Young men who are assigned to the two officer ranks (sergeant and 2nd lieutenant) find themselves among a strongly positively selected group of young men at a critical stage in their life, when they are actively choosing whether to pursue higher education.

[46] One should also interpret these post-service estimates with some caution, since they rely on the assumption that there are no important unobservable differences between the treated and control groups that express themselves as they age.

[47] They study cohorts tested between 1970 and 1988. At this time, all men who were fit were required to serve.

Unfortunately, we also see that military service has a number of negative effects, especially on those from less advantaged backgrounds. They perform worse in the labor market and are more engaged in crime, despite being incapacitated while conscripted. How does service generate such negative effects? Could negative peer influence be part of the explanation?

To study this potential mechanism, we take our IV sample and group those who actually serve into "units". These are not the actual units (or platoons) in which they serve, as we do not have access to data with platoon identifiers.[48] Instead, our units are created by grouping men by test year, regiment and rank. There are 114 regiments in our data and four ranks – private, corporal, sergeant and 2nd lieutenant. We drop units with less than 10 members to allow for more accurate measures of average peer characteristics within a unit – namely pre-service crime. The median unit size is 142; the mean is 162 and the maximum is 589. Although, these are not a conscript's true platoon mates, this is the pool from which platoon mates are drawn.

For each conscript, we then calculate the leave-out average pre-service crime rate of all other men in his unit; that is, this average is based solely on the pre-service crime of each young man's peers, excluding himself. Figure 5 shows the distribution of the leave-out average pre-service number of crimes across units, reported separately by rank. This figure clearly shows that pre-service crime is concentrated among units of peers from the lowest ranks. The entire distribution for privates, and to a lesser extent corporals, is markedly shifted to the right. Men from low SES backgrounds tend also to be concentrated in these units; 29 percent of privates have fathers with nine or less years of schooling, compared to 19 percent for corporals, and 12 percent for sergeants and 2nd lieutenants. Thus, one unintentional side-effect of the recruitment and placement process is that high crime, low SES men are concentrated together into smaller units, with intensive exposure over a long period of time.

[48] Platoons were typically comprised of 20 privates and corporals, one sergeant, and one 2nd lieutenant.

But could this exposure then lead to peer effects in crime? In Table 12, we estimate potential peer effects by regressing own post-service crime (at the ages 23-30) on the leave-out mean pre-service crime rate in one's unit. We control for birth month, birth year, municipality, and test year by test office fixed effects, as well as the full set of test day and pre-test day characteristics, including own pre-service crime. The first specification looks at only the baseline relationship between peer criminal history and an individual's post-service crime. The remaining three columns interact peer criminal history with whether the individual has a low educated father, has a criminal history themselves, and is a private or corporal, respectively. Columns (1) – (4) present the extensive margin, while Columns (5) – (8) present the intensive margin. The estimated peer effect for men from advantaged backgrounds (high educated fathers or no criminal history) or with higher military ranks is zero; exposure to peers with a criminal history does not increase post-service crime for individuals with a low risk of crime to start with. In contrast, the results in Table 12 are indicative of strong peer influences for individuals from disadvantaged backgrounds; increased exposure to peers with a criminal history prior to service is associated with higher post-service crime for conscripts assigned to lower military ranks, from lower SES households, and especially, with a criminal history prior to service themselves. For example, having a criminal record prior to service increases the likelihood of committing a crime post-service by almost nine percentage points; evaluated at the mean, exposure to peers with a criminal history further increases the likelihood of post service crime by an additional three percentage points. In this way, peer effects appear to be reinforcing in nature – exposure to peers with a criminal history reinforces the criminal path that individuals are already on.[49] Such non-linear peer effects imply that how conscripts are allocated to a unit

[49] This is consistent with the findings of reinforcing peer effects in juvenile correctional facilities by Bayer, Hjalmarsson, and Pozen (2009), though those were crime-specific in nature.

can affect post-service crime. Thus, one way to limit the potential negative (unintended) effects of military service may be to not group all “bad apples” together.[50]

Taken together, the concentration of high crime men into the same units of service and potential strong peer effects make negative peer effects one plausible mechanism behind the negative effects that we find for low SES men. Quantifying this effect, however, is quite difficult, since we lack a well-defined measure of the counterfactual peer groups that these young men would have faced if they had not been placed in service. Not only is the composition of the counterfactual peer group difficult to identify, but also the intensity of exposure – military service not only changes the composition of peers but also the intensity of the peer interactions. In this way, even if the average peer characteristics are not that different, peer effects of service could be quite strong since the nature of the interactions has changed.

Of course, there are other potential mechanisms that could also be at work. In particular, we would like to point out the fact that our cohorts of men were leaving service and entering into an environment with rather high unemployment among young adults. One conjecture is that among low skilled workers, those who do not undertake military service are able to establish themselves on the labor market more quickly. As we saw in Table 9, the marginal conscript experiences more unemployment as a young adult.

An additional explanation often put forth in the literature is that a desensitization to violence and weapons can exacerbate post-service crime. However, we cannot use our empirical framework to directly study this channel since assignment to a high service officiator does not impact whether one is in a combat position (see Table 4). Though such an explanation would seem more feasible during wartime enlistment, we cannot conclusively rule it out as playing a role in the current context.

[50] However, as demonstrated by Carrell, Sacerdote, and West (2013), caution should be exercised when trying to optimally design peer groups based on reduced form peer effects. They in fact found perverse effects of the ‘optimal’ assignments to squadrons in the United States Airforce Academy.

7. Conclusion

With the end of the Cold War, numerous countries in Europe abolished mandatory conscription. With no imminent military threat, and with the security of NATO or EU membership, it became hard for politicians to both justify the financial costs of such a large-scale national policy and to convince voters of the need for it on civic grounds alone (Bieri, 2015). In recent years, the debate has about-faced, with many countries considering a re-instatement of mandatory conscription in some form. Perhaps not surprisingly, with the annexation of Crimea by Russia, both Lithuania and the Ukraine have already reinstated the draft. However, this conversation is also happening in countries in Western Europe – namely France, Italy, the UK and Sweden (Bieri, 2015). While one should clearly debate these issues with respect to the direct costs and likely need for and competence of such a military, one must also consider the potential indirect costs associated with mandatory conscription.

Using an instrumental variable approach that takes advantage of exogenous variation in the likelihood of service due to randomly assigned draft board officiators, we show that the potential indirect costs of mandatory conscription may indeed be high. Our instrumental variable analysis finds that conscription significantly increases post-service crimes from ages 23 to 30 across a number of crime categories, including violent crimes and thefts. In addition, these detrimental effects of service are driven by relatively 'high risk' populations with respect to future crime (i.e. those with a criminal history prior to service and/or from low socioeconomic status backgrounds). We provide evidence that grouping high crime, low SES individuals together in an environment with high intensity peer exposure may be one feasible explanation for the negative effects of service for these high risk populations – i.e. reinforcing peer effects. However, we cannot rule out a weaker labor market attachment or desensitization to weapons and violence as explanations of the results. Unfortunately, regardless of the

mechanism, these results contradict the idea that military service may be a way to straighten out troubled youths and build skills that are marketable in the post-service labor market.

On a brighter note, we find that mandatory conscription may improve the health of all young men and can have other positive post-service effects for populations at low-risk for crime. In the instrumental variable analysis, there is little post-service effect on crime and there are actually improved labor market outcomes for individuals from non-criminal backgrounds and better educated families. In addition, we provide the first empirical evidence that military service can incapacitate criminal behavior during the time of service. Unfortunately, our instrumental variable analysis finds that any long term crime reducing effects of incapacitation are more than offset by other ways in which the military negatively affects crime.

Taken together, the results of our analysis indicate mandatory military conscription does have a significant impact on the life course of young men, and that this impact is quite heterogeneous, such that it may reinforce already existing inequalities in the likelihood of future success. These non-monetary costs (and/or benefits) should be taken into account when deciding whether to reinstate or abolish mandatory conscription or when devising the system through which conscription occurs (e.g. lottery, testing, etc.). Who are the average and marginal conscripts? How will conscription affect these individuals?

References

Aizer, Anna and Joe Doyle (2015) "Juvenile Incarceration, Human Capital and Future Crime: Evidence from Randomly-assigned Judges," *Quarterly Journal of Economics* 130(2), 759-803.

Albaek, Karsten, Søren Leth-Petersen, Daniel le Maire and Torben Tranaes (forthcoming) "Does Peacetime Military Service Affect Crime?" *Scandinavian Journal of Economics*.

Albrecht, James W., Per-Anders Edin, Marianne Sundström and Susan B. Vroman (1999) "Career Interruptions and Subsequent Earnings: A Reexamination Using Swedish Data," *Journal of Human Resources* 34(2), 294-311.

Anderson, D. Mark and Daniel Rees (2015) "Deployments, Combat Exposure, and Crime," *Journal of Law and Economics* 58, 235-267.

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

Angrist, Joshua D. and Stacey H. Chen (2011) "Schooling and the Vietnam Era GI Bill: Evidence from the Draft Lottery," *American Economic Journal: Applied Economics* 3(2), 96-118.

Angrist, Joshua D., Stacey H. Chen and Brigham R. Frandsen (2010) "Did Vietnam Veterans Get Sicker in the 1990s? The Complicated Effects of Military Service on Self-Reported Health," *The Journal of Public Economics* 94(11-12), 824-837.

Angrist, Joshua D., Stacey H. Chen and Jae Song (2011) "Long-term Consequences of Vietnam-Era Conscription: New Estimates Using Social Security Data," *American Economic Review: Papers and Proceedings* 101(3), 334–338.

Autor, David, Mark Duggan and David Lyle (2011) "Battle Scars: the Puzzling Decline in Employment and Rise in Disability Receipt among Vietnam-Era Veterans," *American Economic Review: Papers and Proceedings* 101(3), 339-349.

Bayer, Patrick, Randi Hjalmarsson and David Pozen (2009) "Building Criminal Capital Behind Bars: Peer Effects in Juvenile Corrections," *Quarterly Journal of Economics* 124(1), 105-147.

Becker, Gary (1968) "Crime and Punishment: An Economic Approach," *The Journal of Political Economy* 76(2), 169-217.

Beckerman, Adela and Leonard Fontana (1989) "Vietnam Veterans and the Criminal Justice System: A Selected Review," *Criminal Justice and Behavior* 16 (4), 412-428.

Bedard, Kelly and Olivier Deschênes (2006) "The Long-Term Impact of Military Service on Health Outcomes: Evidence from World War II and Korean War Veterans," *The American Economic Review* 96(1), 176-194.

Bieri, Matthias (2015) “Military Conscription in Europe: New Relevance,” *CSS Analyses in Security Policy*, Center for Security Studies, ETH Zurich.

Bingley, Paul, Petter Lundborg and Stéphanie Vincent Lyk-Jensen (2014) “Opportunity Cost and the Incidence of a Draft Lottery,” IZA DP No. 8057.

Böhlmark, Anders and Matthew J. Lindquist (2006) “Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden,” *Journal of Labor Economics* 24(4), 879-896.

Card, David and Ana Rute Cardoso (2012) “Can Compulsory Military Service Raise Civilian Wages? Evidence from the Peacetime Draft in Portugal,” *American Economic Journal: Applied Economics* 4(4), 57-93.

Carlsson, Magnus, Gordon Dahl, Björn Öckert and Dan-Olof Rooth (2015) “The Effect of Schooling on Cognitive Skills,” *Review of Economics and Statistics* 97 (3), 533-547.

Dobkin, Carlos and Reza Shabani (2009). “The Health Effects of Military Service: Evidence from the Vietnam Draft,” *Economic Inquiry* 47(1), 69-80.

Carrell, Scott, Bruce Sacerdote and James West (2013) “From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation,” *Econometrica* 81(3), 855-882.

Doyle Jr., Joseph J. (2008) “Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care,” *Journal of Political* Economy 116(4), 746-770.

Galiani, Sebastian, Martín A. Rossi, and E. Schargrodsky (2011): “The Effects of Peacetime and Wartime Conscription on Criminal Activity," *American Economic Journal: Applied Economics* 3(2), 119-136.

Grenet, J., R. Hart, and Ernesto Roberts (2011) “Above and Beyond the Call: Long-term Real Earnings Effects of British Male Military Conscription in the Post-War Years,” *Labour Economics* 18(2), 194-204.

Grossman, Dave (1995) *On Killing. The Psychological Cost of Learning to Kill in War and Society* (Boston: Little, Brown).

Grönqvist, Erik and Erik Lindqvist (2016) “The Making of a Manager: Evidence from Military Officer Training,” *Journal of Labor Economics* 34(4).

Hanes, Niklas, Erik Norlin and Magnus Sjöström (2010) “The Civil Returns of Military Training: A Study of Young Men in Sweden,” *Defense and Peace Economics* 21(5), 547-565.

Jacob, Brian and Lars Lefgren (2003) “Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime,” *American Economic Review* 93(5), 1560-1577.

Johnston David W., Michael A. Shield, Peter Siminski (forthcoming) “Long-Term Health Effects of Vietnam-Era Military Service: A Quasi-Experiment using Australian Conscription Lotteries,” *Journal of Health Economics*.

Kling, Jeffrey (2006) “Incarceration Length, Employment, and Earnings,” *American Economic Review* 96(3), 863-876.

Lindo, Jason M. and Charles Stoecker (2014) “Drawn into Violence: Evidence on ‘What Makes a Criminal’ from the Vietnam Draft Lotteries,” *Economic Inquiry* 52(1), 239-258.

Luallen, Jeremy (2006) “School's Out...Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes,” *Journal of Urban Economics* 59(1), 75-103.

Maurin, Eric and Theodora Xenogiani (2007) “Demand for Education and Labor Market Outcomes. Lessons from the Abolition of Compulsory Conscription in France,” *Journal of Human Resources* 42(4), 795-819.

Mueller-Smith, Michael (2015) “The Criminal and Labor Market Impacts of Incarceration,” unpublished manuscript, University of Michigan.

Resnick, Heidi S., David W. Foy, Clyde P. Donahoe and Eric N. Miller (1989) “Antisocial Behavior and Post-traumatic Stress Disorder in Vietnam Veterans,” *Journal of Clinical Psychology* 45(6), 860-866.

Rohlfs, Chris (2010) “Does combat exposure make you a more violent or criminal person? Evidence from the Vietnam draft,” *Journal of Human Resources* 45(2), 271-300.

Siminski, Peter (2013) “Employment Effects of Army Service and Veterans Compensation: Evidence from the Australian Vietnam-Era Conscription Lotteries,” *Review of Economics and Statistics* 95(1), 87-97.

Siminski, Peter, Simon Ville, and Alexander Paull (2016). “Does the Military Train Men to Be Violent Criminals? New Evidence from Australia’s Conscription Lotteries,” *Journal of Population Economics* 29(1), 197-218.

Yesavage, Jerome (1983) “Differential Effects of Vietnam Combat Experiences vs Criminality on Dangerous Behavior by Vietnam Veterans with Schizophrenia,” *The Journal of Nervous and Mental Disease* 171(6), 382-384.

Figure 1. Share of Sweden born Males Who Were Tested, Share Assigned to a Service Category, and Share Who Served in the Military by Birth Cohort.

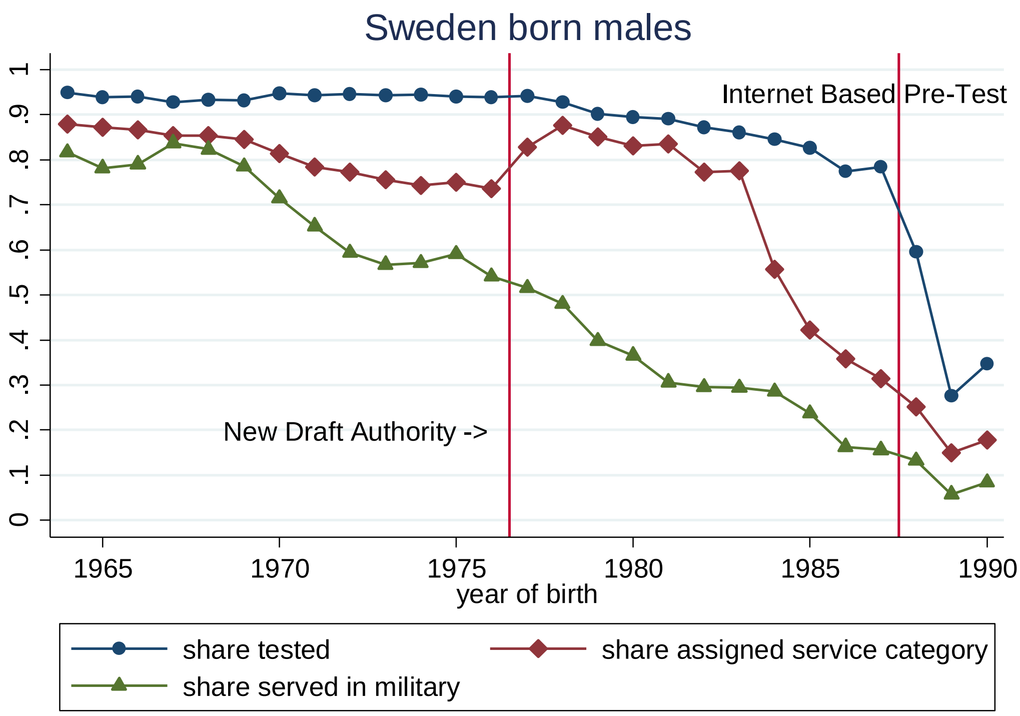
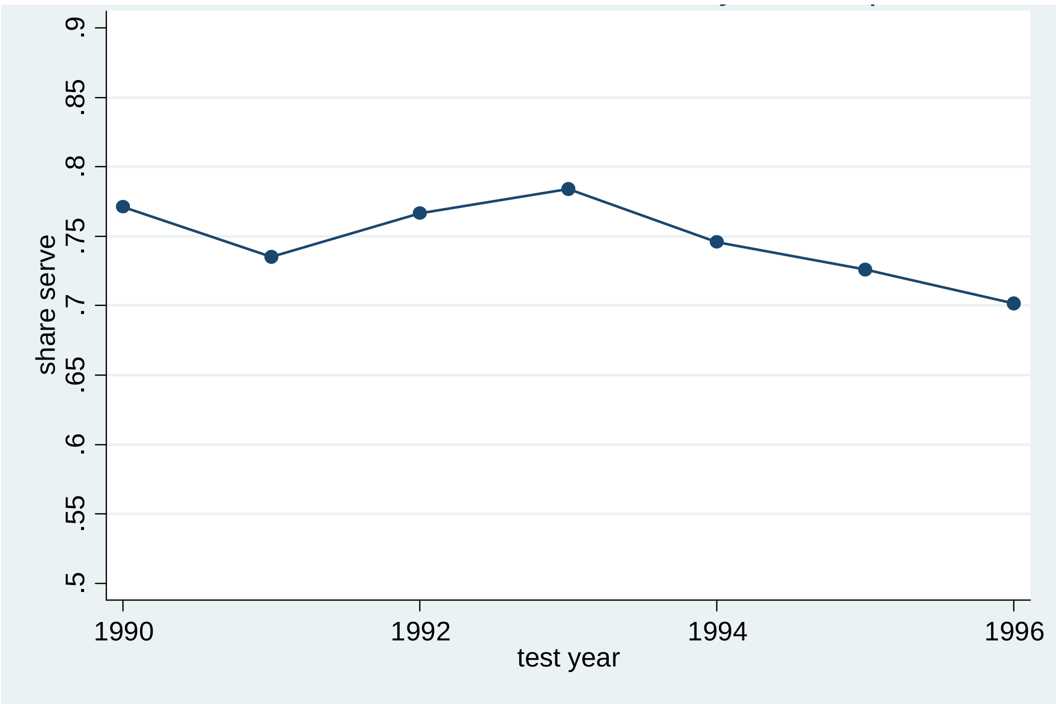


Figure 2. Share Serve from 1990 to 1996 Test Cohorts

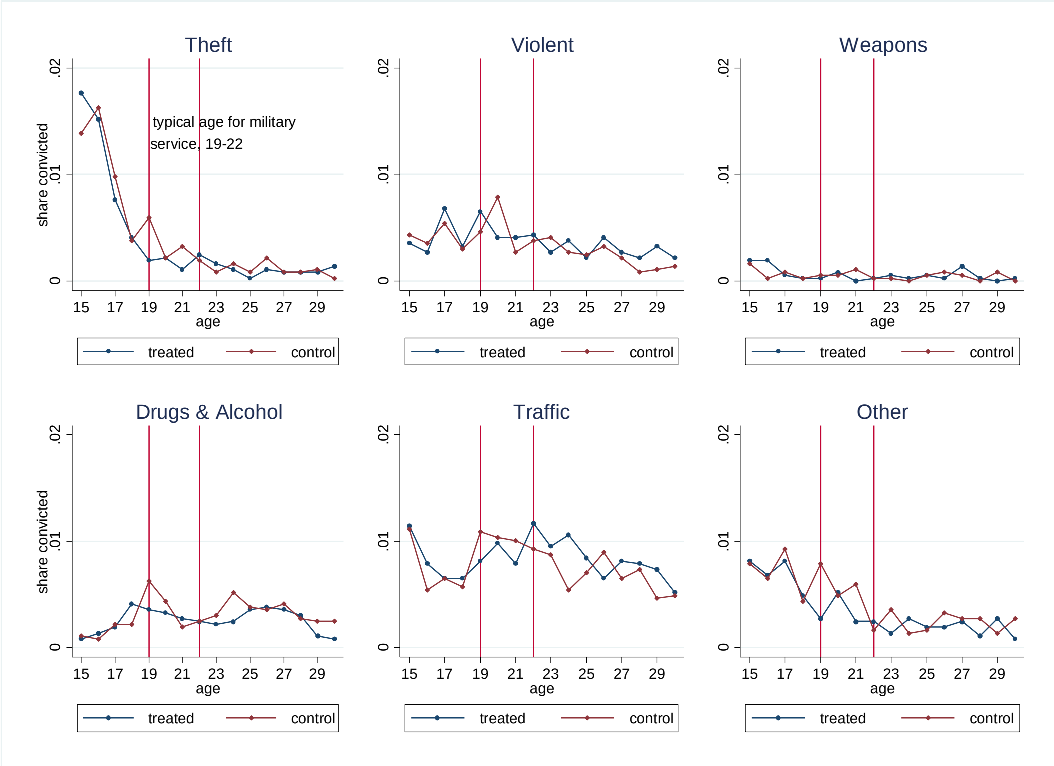


Note – This graph is based on the sample used in the instrumental variable analysis. That is, it is the share that serve of the sample of non-immigrant males who tested in 1990 to 1996 and were eligible to serve. For more details on the sample criteria, see Appendix Table 1.

Figure 3. Crime Trends by Age for Treated and Controls in our Matched Sample, Extensive Margin.

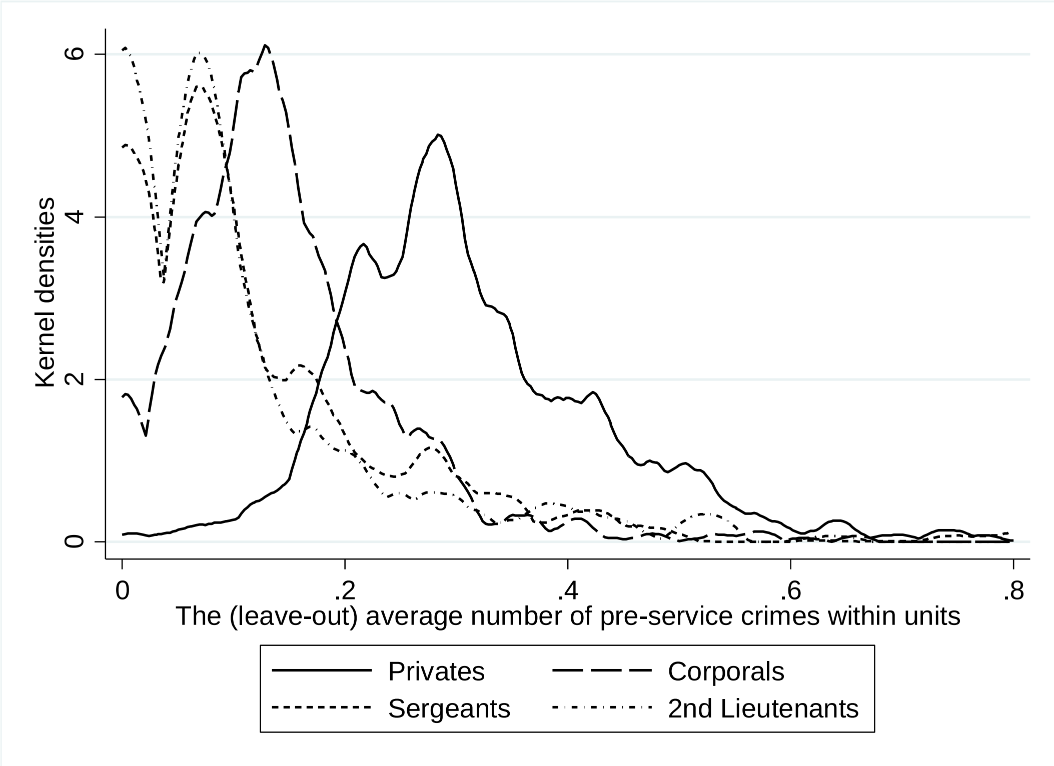
Note – This graph is based on our matched sample, which includes men tested between 1997 and 2001.

Figure 4. Trends in Crime Type by Age for Treated and Controls in our Matched Sample, Extensive Margin.



Note – This graph is based on our matched sample, which includes men tested between 1997 and 2001.

Figure 5. Kernel Densities Over Distribution of the “Leave-Out” Mean Number of Pre-Service Crimes in Each Unit by Rank.



Note – This figure is based on men in our IV sample who actually serve in the military.

Table 1. Summary Statistics: Overall and By Service and High Service Officiator Assignment

			Service =		**High Service Officiator =**	
	All Individuals *N = 168,818*		**1** *N=126,550*	**0** *N =42,268*	**1** *N=85,138*	**0** *N=83,668*
Variable	Mean	SD	Mean	Mean	Mean	Mean
military service (tax records)	0.75	0.43	1.00	0.00	0.83	0.67
high service rate officiator	0.50	0.50	0.56	0.35		
test_year	1992.76	1.98	1992.71	1992.88	1993.04	1992.47
birth_year	1974.56	2.03	1974.55	1974.62	1974.86	1974.27
Pre-test day controls						
Any weapons < 18?	0.01	0.08	0.006	0.008	0.008	0.005
Any violent < 18?	0.02	0.14	0.017	0.023	0.020	0.017
Any traffic < 18?	0.06	0.23	0.054	0.061	0.052	0.059
Any theft < 18?	0.06	0.24	0.057	0.065	0.065	0.054
Any other < 18?	0.04	0.19	0.036	0.044	0.040	0.037
Any drugs< 18	0.00	0.06	0.003	0.004	0.003	0.004
schooling father	11.14	2.65	11.16	11.08	11.33	10.95
schooling mother	11.55	2.35	11.57	11.49	11.71	11.38
income father	12.32	0.40	12.32	12.30	12.33	12.30
income mother	11.87	0.39	11.87	11.85	11.90	11.83
2 year school?	0.18	0.39	0.18	0.19	0.16	0.20
3 year school?	0.76	0.43	0.77	0.72	0.77	0.74
Test day controls						
Height	179.75	6.44	179.77	179.66	179.74	179.75
Weight	71.59	10.53	71.65	71.40	71.55	71.63
BMI	22.14	2.89	22.15	22.10	22.13	22.15
Ability Score	5.17	1.84	5.22	5.05	5.22	5.13
Physical Capacity Score	6.26	1.45	6.35	5.98	6.29	6.22
psychological test score	5.35	1.53	5.49	4.92	5.42	5.28
Selected Outcome variables						
Any crimes 23-30?	0.10	0.31	0.10	0.11	0.11	0.10
# crimes 23-30	0.31	2.33	0.28	0.40	0.33	0.29
2 or more crimes 23-30?	0.04	0.20	0.04	0.05	0.04	0.04
5 or more crimes 23-30?	0.01	0.11	0.01	0.02	0.01	0.01
Any weapons 23-30?	0.01	0.07	0.01	0.01	0.01	0.00
Any violent 23-30?	0.02	0.15	0.02	0.03	0.03	0.02
Any traffic 23-30?	0.06	0.25	0.06	0.07	0.07	0.06
Any theft 23-30?	0.01	0.11	0.01	0.02	0.01	0.01
Any other 23-30?	0.02	0.15	0.02	0.03	0.03	0.02
And drugs 23-30?	0.02	0.13	0.01	0.02	0.02	0.01
More than 12 years school?	0.42	0.49	0.42	0.41	0.43	0.41
(log) income_30_34	12.38	0.78	12.40	12.33	12.39	12.38
Any welfare 23-34?	0.13	0.34	0.12	0.15	0.13	0.13
Any sick days 23-34?	0.33	0.47	0.33	0.35	0.32	0.34
Any disability pension 23-34?	0.01	0.11	0.01	0.02	0.01	0.01
Any unemploy. benefits 23-34?	0.46	0.50	0.46	0.45	0.46	0.46

Note - Missing observations are replaced with sample means (for the control variables only).

Table 2. Test of Random Assignment to Officiators

	(1)	(2)	(3)	(4)	(5)
	Dependant Variable = High Service officiator				unexp
# (of 24) coefficients significant 5% (including missing dummies)	12	6	3	1	0
# (of 18) coefficients significant 5% (excluding missing dummies)	7	5	2	1	0
p-value observable controls	0	0.08	0.25	0.14	0.12
county FE	no	yes	yes	yes	no
test year FE	no	no	yes	yes	no
county x test year FE	no	no	no	yes	no
Pre Characteristics					
Any weapons < 18?	0.08582**	0.00533	-0.00962	0.00632	0.00641
	[0.03270]	[0.00751]	[0.00932]	[0.00743]	[0.00705]
Any violent < 18?	0.02887	0.01503**	0.00113	0.00036	0.00039
	[0.02125]	[0.00616]	[0.00439]	[0.00376]	[0.00568]
Any traffic < 18?	-0.03784***	0.00436	0.00511	0.00590*	0.0059
	[0.00786]	[0.00452]	[0.00413]	[0.00342]	[0.00477]
Any theft < 18?	0.06094***	0.00421	0.00279	0.00052	0.00052
	[0.00811]	[0.00311]	[0.00280]	[0.00216]	[0.00407]
Any other < 18?	0.01322*	-0.00465	-0.00268	-0.00379	-0.00379
	[0.00703]	[0.00388]	[0.00344]	[0.00307]	[0.00326]
Any drugs< 18	-0.02941	0.00813	0.00782	0.00941	0.00939
	[0.02263]	[0.00914]	[0.00938]	[0.00864]	[0.01219]
scholling father	0.00791***	0.00043*	-0.00039	0.00009	0.00009
	[0.00119]	[0.00025]	[0.00027]	[0.00019]	[0.00050]
schooling mother	0.00320***	0.00217***	0.00051	-0.00002	-0.00002
	[0.00118]	[0.00067]	[0.00040]	[0.00034]	[0.00047]
income father	0.00048	-0.01011***	-0.00165	-0.00274*	-0.00267
	[0.01868]	[0.00345]	[0.00177]	[0.00156]	[0.00454]
income mother	0.08930***	0.0025	-0.0005	0.00154	0.00159
	[0.01531]	[0.00214]	[0.00198]	[0.00149]	[0.00419]
2 year school?	-0.05725*	-0.04499***	0.00880**	0.00403*	0.00386
	[0.03079]	[0.01370]	[0.00363]	[0.00225]	[0.01091]
3 year school?	-0.01635	0.01923*	0.01265***	0.00482*	0.00479
	[0.02547]	[0.01093]	[0.00462]	[0.00267]	[0.00969]
Test Day Characteristics					
Height	0.00014	-0.00012	-0.00038	-0.0009	-0.00089
	[0.00218]	[0.00075]	[0.00077]	[0.00077]	[0.00089]
Weight	-0.00118	0.00011	0.0006	0.00118	0.00117
	[0.00260]	[0.00094]	[0.00098]	[0.00095]	[0.00109]
BMI	0.00383	0.00083	-0.00204	-0.00431	-0.00427
	[0.00765]	[0.00333]	[0.00329]	[0.00314]	[0.00354]
ability test score	-0.00574**	-0.00456***	-0.00197	-0.00148**	-0.00147
	[0.00267]	[0.00143]	[0.00120]	[0.00062]	[0.00110]
physical capacity score	0.00052	-0.00151	0.00261	0.00101*	0.00095
	[0.01022]	[0.00251]	[0.00237]	[0.00053]	[0.00296]
psychological test score	0.01112	0.00154	0.00183	0.00111*	0.0011
	[0.00733]	[0.00154]	[0.00157]	[0.00063]	[0.00108]
Observations	168806	168806	168806	168806	168806
R-squared	0.02	0.62	0.66	0.75	0

Robust standard errors (clustered at officiator) .The dependent variable in column (5) is the residual from a regression of assignment to a high service rate officiator on test year * county fixed effects, i.e. the variation in officiator assignment not explained by geographical and test year fixed effects.

Table 3. First Stage Regressions of Military Service on Assignment to High Service Rate Officiator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable = military service							
	Full Sample				Prior Criminal History	No History	Father <=9 yearsschool	Father >9 years school
High Service Rate Officiator	0.15513***	0.10229***	0.07933***	0.07463***	0.09703***	0.07100***	0.08854***	0.06921***
	[0.01999]	[0.01166]	[0.01398]	[0.01349]	[0.01850]	[0.01340]	[0.01439]	[0.01380]
F-Statistic	60	77	32	31	28	28	38	25
county FE	no	yes	yes	yes	yes	yes	yes	yes
test year FE	no	yes	yes	yes	yes	yes	yes	yes
county x test year FE	no	no	yes	yes	yes	yes	yes	yes
test office FE	no	no	no	yes	yes	yes	yes	yes
test office x test year FE	no	no	no	yes	yes	yes	yes	yes
pre-test day characteristics	no	no	no	yes	yes	yes	yes	yes
test day variables	no	no	no	yes	yes	yes	yes	yes
Observations	168806	168806	168806	168806	22590	146216	43904	124902
R-squared	0.03	0.06	0.07	0.09	0.1	0.09	0.1	0.09

Robust standard errors, clustered by officiator. *** significant at 1%, ** significant at 5%, * significant at 10%

Table 4. Do High Service Officiators Affect Type of Service?

	Dependent Variable:							
	military service	army	navy	airforce	private	corporal	sergeant or 2nd lieutenant	combatant
High Service Rate Officiator	0.07933***	-0.00817	0.01068*	-0.00399	-0.00704	0.00944	-0.0023	-0.01415
	[0.01398]	[0.00920]	[0.00591]	[0.00746]	[0.01466]	[0.01060]	[0.00622]	[0.01024]
F-Statistic	32	1	3	0	0	1	0	2
county FE	yes	yes	yes	yes	yes	yes	yes	yes
test year FE	yes	yes	yes	yes	yes	yes	yes	yes
county x test year FE	yes	yes	yes	yes	yes	yes	yes	yes
Observations	168806	165928	165928	165928	168806	168806	168806	152158
R-squared	0.07	0.08	0.09	0.05	0.01	0	0.01	0.01

Robust standard errors in brackets, clustered on officiator

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5. OLS, Reduced Form, and IV Baseline Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		OLS		Reduced Form		IV (reduced form/first stage)	
Panel A: Dependent Variable = Any Crimes from Age 23-30 (Mean = 0.10)							
military_service	-0.01233***	-0.01661***	-0.00485**	0.00707*	0.00636**	0.08908*	0.08448*
	[0.00283]	[0.00277]	[0.00203]	[0.00358]	[0.00293]	[0.04779]	[0.04505]
Panel B: Dependent Variable = # Crimes from Age 23-30 (Mean = 0.31)							
military_service	-0.12176***	-0.13289***	-0.04827***	0.04431**	0.04220**	0.55856**	0.56039**
	[0.01775]	[0.01965]	[0.01648]	[0.02080]	[0.01801]	[0.25192]	[0.23286]
Panel C: Dependent Variable = # Crimes from Age 23-30 >=2 (Mean = 0.042)							
military_service	-0.01106***	-0.01294***	-0.00444***	0.0032	0.00299	0.04028	0.03976
	[0.00151]	[0.00151]	[0.00108]	[0.00238]	[0.00210]	[0.02713]	[0.02541]
Panel D: Dependent Variable = # Crimes from Age 23-30 >=5 (Mean = 0.013)							
military_service	-0.00566***	-0.00610***	-0.00197**	0.0016	0.00136	0.02014	0.01803
	[0.00081]	[0.00084]	[0.00074]	[0.00121]	[0.00105]	[0.01447]	[0.01403]
First Stage F-Statistic						32	32
First Stage Coefficient						0.07933	0.0753
County x Test Year FE	no	yes	yes	yes	yes	yes	yes
Test office Fixed Effects	no	no	yes	no	yes	no	yes
Pre-test and Test day Controls	no	no	yes	no	yes	no	yes

Columns (1) – (3) present the results of regressing crime from 23-30 (at the extensive margin in Panel A and various intensive margin measures in Panels B-D) on military service and the indicated set of controls. Note that county x test year fixed effects implies the inclusion of county dummies, test year dummies, and county by test year dummies. For the ease of presentation, just the coefficient on military service is reported. Columns (4) and (5) present the reduced form, i.e. regressions of crime from 23-30 on assignment to a high service officiator. Finally, columns (6) and (7) instrument for military service with assignment to a high service rate officiator. Robust standard errors, clustered by county in columns (1) - (3) and officiator in columns (4)-(7).

*** significant 1%, ** significant 5%, * significant 10%. N= 168806

Table 6. Instrumental Variable Estimates of the Effect of Service on Post-Conscription Crime, by Crime Type

	Dependent Variable:						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any Crime	Weaspons	Violent	Traffic	Theft	Other	Drugs/Alc
Panel A. extensive margin 23-30							
(a) Baseline	0.08908*	0.01650**	0.04934***	0.02943	0.01967	0.01886	0.01664
	[0.04779]	[0.00823]	[0.01609]	[0.03362]	[0.01322]	[0.02053]	[0.01527]
(b) +test office, pre testday and test day controls	0.08448*	0.01633**	0.04877***	0.02425	0.01936	0.02003	0.01642
	[0.04505]	[0.00818]	[0.01727]	[0.03314]	[0.01272]	[0.01964]	[0.01333]
Mean Dependent Variable	*0.1*	*0.0058*	*0.026*	*0.064*	*0.013*	*0.025*	*0.016*
Panel B. # crimes 23-30							
(a) Baseline	0.55856**	0.01272	0.10193**	0.05209	0.17040***	0.10175*	0.11969**
	[0.25192]	[0.02072]	[0.04728]	[0.15086]	[0.05596]	[0.05626]	[0.05645]
(b) +test office, pre testday and test day controls	0.56039**	0.00931	0.09217*	0.05008	0.17384***	0.11250**	0.12249**
	[0.23286]	[0.02165]	[0.04759]	[0.15053]	[0.05633]	[0.05588]	[0.05303]
Mean Dependent Variable	*0.31*	*0.01*	*0.046*	*0.13*	*0.038*	*0.046*	*0.047*
Panel C. 2 or more crimes 23-30?							
(a) Baseline	0.04028	0.00445	0.00658	0.01022	0.01830***	0.01161	0.02368**
	[0.02713]	[0.00419]	[0.01033]	[0.01648]	[0.00708]	[0.01129]	[0.01056]
(b) +test office, pre testday and test day controls	0.03976	0.00377	0.00413	0.01011	0.01861***	0.0111	0.02451**
	[0.02541]	[0.00440]	[0.01075]	[0.01600]	[0.00659]	[0.01053]	[0.01081]
Mean Dependent Variable	*0.042*	*0.0018*	*0.009*	*0.02*	*0.0051*	*0.0076*	*0.0079*

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*. N = 168,806. The baseline regressions control for test year, county and test year x county dummies. These controls are also included in all other regressions. All specifications instrument for military service with assignment to a high service rate officiator. Just the coefficient on military service is presented. The first-stage F statistic equals 32 both with and without the full set of controls.

Table 7. Heterogeneity: By Criminal History and Father Schooling

	Dependent Variable: At Least One Conviction from age 23-30 of						
	Any Crime	Weaspons	Violent	Traffic	Theft	Other	Drugs/Alc
Panel A: Heterogeneity By Criminal History							
Sample: At least one crime < 18 (n = 22590)	0.22203**	0.08171	0.10332*	0.09591	0.12952**	0.09816	0.07183
	[0.10382]	[0.05014]	[0.05762]	[0.07425]	[0.06512]	[0.08492]	[0.06502]
Mean Dep Variable	*0.24*	*0.023*	*0.072*	*0.15*	*0.048*	*0.071*	*0.06*
Sample: No crimes < 18 (n=146216)	0.05617	0.00243	0.03660**	0.0135	-0.00372	0.00058	0.00336
	[0.03987]	[0.00656]	[0.01497]	[0.03232]	[0.01220]	[0.01522]	[0.01111]
Mean Dep Variable	*0.082*	*0.0028*	*0.015*	*0.052*	*0.0076*	*0.017*	*0.0093*
Panel B: Heterogeneity by Father Schooling							
Sample: Father Schooling <= 9 years (n = 43904)	0.10682*	0.02667*	0.08606**	0.02519	0.04018	0.03325	0.05055**
	[0.06044]	[0.01458]	[0.03580]	[0.04161]	[0.02724]	[0.03382]	[0.02493]
Mean Dep Variable	0.12	0.0074	0.03	0.074	0.017	0.032	0.02
Sample: Father Schooling > 9 years (n=124902)	0.08244	0.01126	0.03234	0.03288	0.00933	0.01203	0.0005
	[0.05576]	[0.01077]	[0.02008]	[0.03871]	[0.01428]	[0.02247]	[0.02000]
Mean Dep Variable	0.098	0.0049	0.02	0.061	0.012	0.022	0.014

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*. Military Service is instrumented for with the high service rate officiator dummy. F-statistic on first stage regressions: 27 for sample with criminal history, 30 for sample without criminal history, 41 for low education fathers and 27 for high education fathers. Each regression controls for test year dummies, county dummies and test year x county dummies.

Table 8. Falsification Test – Alternative Test of Random Assignment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent Variable: Any Conviction Age 23-30?									
	all	criminal history	no criminal history	father school <= 9	father school > 9	all	criminal history	no criminal history	father school <= 9	father school > 9
Panel A: Main Analysis Sample - Reduced Form										
High Service Officiator	0.00694*	0.02200**	0.00422	0.00949*	0.00612	0.00629*	0.02139**	0.00409	0.00798	0.00568
	[0.00400]	[0.00975]	[0.00313]	[0.00551]	[0.00434]	[0.00355]	[0.00873]	[0.00318]	[0.00508]	[0.00385]
Mean Dependent variable	0.1	0.25	0.082	0.12	0.098	0.1	0.25	0.082	0.12	0.098
Observations	168881	22610	146271	43923	124958	168881	22610	146271	43923	124958
Panel B: Falsification Sample - Omitted Health Categories - Reduced Form										
High Service Officiator	0.00864	0.00415	0.00349	0.00631	0.00764	-0.00148	-0.01276	0.00115	-0.00487	-0.00141
	[0.00860]	[0.01485]	[0.00629]	[0.01303]	[0.00930]	[0.00592]	[0.01427]	[0.00623]	[0.01117]	[0.00716]
Mean Dependent Variable	0.2	0.4	0.13	0.23	0.19	0.2	0.4	0.13	0.23	0.19
Observations	32804	8571	24233	10593	22211	32804	8571	24233	10593	22211
County x Test Year fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Test office Fixed Effects	no	no	no	no	no	yes	yes	yes	yes	yes
Pre-test and Test day Controls	no	no	no	no	no	yes	yes	yes	yes	yes

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*.

Table 9. IV Estimates for Non-Crime Outcomes: Education, Income, and Unemployment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	all	at least 1 pre service crime	no pre service crime	father schooling <= 9 years	father schooling > 9 years	all	at least 1 pre service crime	no pre service crime	father schooling <= 9 years	father schooling > 9 years
	Dep var = schooling > 12 years					*Dep Var = income_30_34*				
military_service_tax_record	-0.05729	0.0034	-0.05884	-0.04076	-0.06957	0.04554	-0.29046	0.12098	-0.29374**	0.20616**
	[0.07003]	[0.10675]	[0.07487]	[0.07045]	[0.09449]	[0.07855]	[0.23359]	[0.07394]	[0.13010]	[0.09475]
Mean Dep Variable	0.42	0.22	0.45	0.27	0.47	12.38	12.25	12.4	12.35	12.4
	Dep Var = any years unemployment benefit 23-34					*Dep Var = # years unemployment benefit 23-34*				
military_service_tax_record	0.019	0.20030**	-0.02611	0.14210*	-0.03593	0.0748	0.26205	0.00124	0.59495*	-0.16777
	[0.06547]	[0.09499]	[0.07547]	[0.07570]	[0.08121]	[0.28268]	[0.54141]	[0.33678]	[0.31594]	[0.35119]
Mean Dep Variable	0.46	0.55	0.45	0.5	0.044	1.55	2.04	1.48	1.79	1.47
	Dep var = any welfare from 23-34					*Dep Var = months welfare 23-34*				
military_service_tax_record	0.0486	0.12019	0.02677	0.10136	0.0241	1.75942*	8.28290**	0.31953	4.98632**	0.22014
	[0.04615]	[0.10392]	[0.04122]	[0.07911]	[0.05258]	[0.99210]	[3.81385]	[0.77772]	[2.27897]	[1.06789]
Mean Dep Variable	0.13	0.25	0.11	0.15	0.12	1.56	4.2	1.15	1.98	1.41
	Dep var = any_disability pension 23-34					*Dep Var = # years with full/partial disability pension*				
military_service_tax_record	-0.02178**	-0.06129**	-0.01564*	-0.02628	-0.0186	-0.10712*	-0.13709	-0.1112**	-0.07816	-0.11637
	[0.01041]	[0.02876]	[0.00946]	[0.02171]	[0.01446]	[0.05570]	[0.16275]	[0.05527]	[0.11669]	[0.07078]
Mean Dep Variable	0.012	0.024	0.01	0.014	0.011	0.057	0.113	0.048	0.067	0.053
	Dep Var = # sick days 23-34									
military_service_tax_record	-59.7342***	-107.6723*	-52.9105***	-44.001	-65.7360***					
	[21.16941]	[55.38455]	[19.99828]	[42.55824]	[24.10553]					
Mean Dep Variable	48.8	83.7	43.5	59.1	45.2					

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*. First stage f-stats: 32 for whole sample, 27 for preservice crime = 1, 30 for no pre service crime, 41 for father <=9 years school, 27 for father >9 years school. Each regression controls for test year dummies, county dummies and test year x county dummies, and instruments for military service with assignment to high service officiator. Samples sizes are N=168806 for the whole sample, 22590 for criminal history, 146216 for no criminal history, 43904 for father schooling<=9, and 124902 for father schooling > 9 samples for all outcomes other than income. For income, about 3500 observations are missing.

Table 10. IV Estimates of a Potential Incapacitation Effect: Crime at 19 and 20 for 18-year Old Testees

	Dependent Variable:						
	Any Crime	Weapons	Violent	Traffic	Theft	Other	Drugs/Alc
Panel A: extensive margin crime at age 19 and 20							
Sample: all 18 year old testees	-0.01972	0.00076	-0.02590*	0.01743	0.00541	-0.00299	-0.01086
(N = 138799)	[0.02825]	[0.00442]	[0.01497]	[0.02093]	[0.01302]	[0.01552]	[0.00801]
Mean	0.075	0.003	0.015	0.028	0.019	0.025	0.005
Sample: 18 year old testees with criminal history	-0.1185	0.02526	-0.05726	-0.10307*	0.06234	0.02823	0.00991
(N = 18558)	[0.11841]	[0.02034]	[0.05448]	[0.06046]	[0.05937]	[0.06804]	[0.02739]
Mean	0.198	0.013	0.056	0.068	0.064	0.072	0.018
Sample: 18 year old testees with no history	-0.00076	-0.00392	-0.02127*	0.04161*	-0.00472	-0.00815	-0.01527**
(N=120241)	[0.02988]	[0.00555]	[0.01115]	[0.02364]	[0.01179]	[0.01448]	[0.00741]
Mean	0.056	0.002	0.009	0.021	0.012	0.018	0.003
Panel B: extensive margin crime at age 23-30							
Sample: all 18 year old testees	0.06037	0.00473	0.04644**	0.00923	0.00219	0.00984	0.00981
(N = 138799)	[0.04448]	[0.01096]	[0.01898]	[0.03285]	[0.01597]	[0.02344]	[0.01773]
Mean	0.103	0.005	0.023	0.063	0.013	0.024	0.016

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*. To isolate incapacitation, this table focuses on the sample who took the test at age 18 (and likely served from ages 19-20) and looks at crime outcomes at age 19 and 20 in Panel A. As individuals do not serve for all of this period, however, these crime outcomes clearly include some pre- and post-service crime. Panel B demonstrates that the same post-service effect is seen for this subample as the whole sample. Military Service is instrumented for with the high service rate officiator dummy. F-statistic on first stage regressions: 26 for all 18 year old testees, 24 for those with a criminal history, and 23 for those with no history. Each regression controls for test year dummies, county dummies and test year x county dummies.

Table 11. Matching Estimates of the Incapacitation Effect of Military Service Using Exact Service Dates and Exact Crime Dates.

	Dependent Variable:						
	Any Crime	Weapons	Violent	Traffic	Theft	Other	Drugs/Alc
Panel A: extensive margin crime during service							
Matched sample	-0.0098***	-0.0003	-0.0022	-0.0019	-0.0022*	-0.0022	-0.0033***
(N = 7366)	[0.0033]	[0.0005]	[0.0015]	[0.0019]	[0.0013]	[0.0017]	[0.0013]
Mean	0.0197	0.0004	0.0039	0.0064	0.0030	0.0049	0.0031
Matched sample with criminal history	-0.0203	0.0029	-0.0087	0.0058	-0.0087	-0.0087	-0.0116
(N = 690)	[0.0240]	[0.0029]	[0.0156]	[0.0110]	[0.0104]	[0.0129]	[0.0108]
Mean	0.0516	0.0015	0.0197	0.0121	0.0046	0.0137	0.0091
Matched sample with no history	-0.0084***	-0.0006	-0.0009	-0.0033*	-0.0015	-0.0018	-.0024*
(N = 6676)	[0.0033]	[0.0005]	[0.0013]	[0.0020]	[0.0014]	[0.0017]	[0.0013]
Mean	0.0165	0.0003	0.0024	0.0058	0.0028	0.0040	0.0025
Panel B: extensive margin crime at age 23-30							
Matched sample	0.0035	0.0008	0.0046	0.0084*	-0.0003	-0.0024	-0.0046*
(N = 7366)	[0.0063]	[0.0012]	[0.0032]	[0.0050]	[0.0018]	[0.0025]	[0.0029]
Mean	0.0787	0.0027	0.0185	0.0477	0.0058	0.0114	0.0160

Standard errors [in brackets] do not account for the fact that the propensity score is estimated. *** 1%, ** 5%, * 10%. Criminal history == 1 if the individual has at least one conviction between the ages of 15 and 17. A 1 to 1 nearest neighbor (with replacement) matching estimator is applied to our matched sample.

Table 12. Potential Peer Effects in Crime.

Dependent variable:	Any crime between ages 23-30				Number of crimes between ages 23-30			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unit pre-service crime	0.029	0.0079	0.024	-0.027	0.053**	-0.037*	0.007	-0.019
	[0.0194]	[0.0191]	[0.0213]	[0.0452]	[0.0250]	[0.0194]	[0.0251]	[0.0323]
Own pre-service crime = 1	0.115***	0.089***	0.115***	0.115***	0.454***	0.2442***	0.455***	0.454***
	[0.0037]	[0.0109]	[0.0037]	[0.0037]	[0.0249]	[0.0470]	[0.0249]	[0.0249]
Father education <= 9 years			-0.002				-0.042**	
			[0.00061]				[0.0205]	
Private or corporal = 1				-0.007				-0.023**
				[0.0050]				[0.0113]
Unit pre-service crime * Own pre-service crime = 1		0.201**				0.748***		
		[0.0809]				[0.1832]		
Unit pre-service crime * Father education <= 9 years			0.022				0.192**	
			[0.0456]				[0.0790]	
Unit pre-service crime * Private or corporal = 1				0.068				0.083*
				[0.0502]				[0.0430]
Mean dependent variable	0.093	0.093	0.093	0.093	0.213	0.213	0.213	0.213
Mean unit crime	0.123	0.123	0.123	0.123	0.262	0.262	0.262	0.262
If pre-service crime = 1 / = 0		0.131/0.122				0.284/0.259		
If father education <= 9 years / > 9 years			0.129/0.121				0.280/0.256	
If private or corporal = 1 / = 0				0.128/0.067				0.274/0.116
Observations	102,085	102,085	102,085	102,085	102,085	102,085	102,085	102,085

Robust standard errors in brackets; *** p<0.01, ** p<0.05, * p<0.1. This utilizes sample of individuals who tested from 1990-1996 (i.e. the IV sample) who served. All specifications include the full set of test day and pre-test day controls as well as birth month, birth year, municipality and test year by test office fixed effects. Columns (3) and (7) exclude controls for father's years of schooling, which is one of our pre-test day controls, instead using the dummy for fathers having nine or less years of education. Note that unit pre-service crime is the leave out mean crime rate for each individual's unit, as defined by test year, rank, and regiment cells.

Appendix Table 1. Sample Restrictions for Baseline Instrumental Variable Analysis

Sample Restriction	Sample Size
Non-immigrant males born 1964-1990	955,454
Exclude if missing test year information	784,055
Excluding if missing county identifier	783,508
drop if test year>=1997 or test year < 1990	231,583
drop testees with officiators who see less than 100 cases in their test year	216,707
drop those missing health group information or assigned to health groups that 'never' serve in that year	174,489
Keep only those who are assigned to service	173,601
drop those with unknown end year of service and those 23 or older in year finish service	168,818

This figure provides sample criteria used in creating the baseline sample used in the high service rate officiator instrumental variable analysis.

Appendix Table 2. Sensitivity Analysis of Baseline Specification to Standard Error Clustering and Alternative Instrument

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	All	All	All	All	All
Panel A: Dependent Variable = Any Crimes from Age 23-30 (Mean = 0.10)						
military_service	0.08908*	0.08908**	0.08908	0.08908*	0.05707**	0.04999**
	[0.04779]	[0.04104]	[0.05673]	[0.04731]	[0.02613]	[0.02356]
Panel B: Dependent Variable = # Crimes from Age 23-30 (Mean = 0.31)						
military_service	0.55856**	0.55856**	0.55856**	0.55856**	0.54830***	0.53779***
	[0.25192]	[0.24409]	[0.28332]	[0.26652]	[0.18881]	[0.18207]
Instrument	high service officiator	high service officiator	high service officiator	high service officiator	leave out annual mean	leave out annual mean
Cluster Unit	officiator	county	2 way: officiator and test year	officiator x test year	officiator	officiator
Number of clusters	67	24	67 and 7	203	67	67
First Stage F-Statistic	32	116	45	39	79	87
county x test year fixed effects	yes	yes	yes	yes	yes	yes
Test office Fixed Effects	no	no	no	no	no	yes
Pre-test and Test day Controls	no	no	no	no	no	yes

Columns (1) - (4) present the results of instrumenting for service with a high service rate officiator dummy while columns (5) and (6) uses the leave out annual mean, i.e. the actual share of testees assigned to an officiator in a given year, excluding the candidate under consideration, who are assigned to service. Standard errors are clustered, and first stage f-statistics calculated accordingly, according to the notes in the table. *** significant 1%, ** significant 5%, * significant 10%. N= 168806

Appendix Table 3. IV Estimates on Post-Conscription Crime by Crime Type, Using the Leave Out Annual Mean as the Instrument

	Dependent Variable:						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any Crime	Weapons	Violent	Traffic	Theft	Other	Drugs/Alc
Panel A. extensive margin 23-30							
Baseline	0.05707**	0.01105*	0.01893*	0.03528*	0.00821	0.01309	0.01213
	[0.02613]	[0.00609]	[0.00991]	[0.01804]	[0.00892]	[0.01271]	[0.00947]
Mean Dependent Variable	*0.1*	*0.0058*	*0.026*	*0.064*	*0.013*	*0.025*	*0.016*
Panel B. # crimes 23-30							
Baseline	0.54830***	0.03250**	0.06415**	0.18665**	0.10810**	0.0332	0.12371***
	[0.18881]	[0.01349]	[0.03124]	[0.09316]	[0.04855]	[0.03612]	[0.03845]
Mean Dependent Variable	*0.31*	*0.01*	*0.046*	*0.13*	*0.038*	*0.046*	*0.047*
Panel C. 2 or more crimes 23-30?							
Baseline	0.02383	0.00687**	0.00452	0.0158	0.00749	0.01904**	0.01308*
	[0.01973]	[0.00293]	[0.00696]	[0.00975]	[0.00621]	[0.00956]	[0.00750]
Mean Dependent Variable	*0.042*	*0.0018*	*0.009*	*0.02*	*0.0051*	*0.0076*	*0.0079*

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*. N = 168,806 (corresponding to 1990-1996 test years). The baseline regressions control just for test year dummies, county dummies and test year x county dummies. These controls are also included in all other regressions. All specifications instrument for military service with the leave out annual mean. Just the coefficient on military service is presented. The first-stage F statistic equals 79.

Appendix Table 4. Heterogeneity: By Criminal History and Father Schooling -- Leave Out Annual Mean as Instrument

	Dependent Variable: At Least One Conviction from age 23-30 of						
	Any Crime	Weapons	Violent	Traffic	Theft	Other	Drugs/Alc
Panel A: Heterogeneity By Criminal History							
Sample: At least one crime < 18 (n = 22590)	0.11406	0.07386*	0.04106	0.10046	0.11596*	0.07369	0.08628
	[0.12719]	[0.04442]	[0.05118]	[0.09298]	[0.06150]	[0.06029]	[0.06037]
Mean Dep Variable	*0.24*	*0.023*	*0.072*	*0.15*	*0.048*	*0.071*	*0.06*
Sample: No crimes < 18 (n=146216)	0.03547	-0.00025	0.01086	0.01834	-0.01262	-0.00044	-0.0038
	[0.02212]	[0.00421]	[0.00947]	[0.01575]	[0.00793]	[0.00982]	[0.00814]
Mean Dep Variable	*0.082*	*0.0028*	*0.015*	*0.052*	*0.0076*	*0.017*	*0.0093*
Panel B: Heterogeneity by Father Schooling							
Sample: Father Schooling <= 9 years (n = 43904)	0.02653	0.02324**	0.04573*	0.02315	0.01523	0.00309	0.04520**
	[0.04295]	[0.01171]	[0.02595]	[0.03982]	[0.01834]	[0.02012]	[0.02094]
Mean Dep Variable	0.12	0.0074	0.03	0.074	0.017	0.032	0.02
Sample: Father Schooling > 9 years (n=124902)	0.07326**	0.00675	0.00911	0.04252**	0.00615	0.01862	-0.00023
	[0.03022]	[0.00662]	[0.01214]	[0.02029]	[0.00783]	[0.01613]	[0.01203]
Mean Dep Variable	0.098	0.0049	0.02	0.061	0.012	0.022	0.014

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*. Military Service is instrumented for with the leave out annual mean. F-statistic on first stage regressions: 47 for sample with criminal history, 80 for sample without criminal history, 63 for low education fathers and 78 for high education fathers. Each regression controls for test year dummies, county dummies and test year x county dummies.

Appendix Table 5. IV Estimates for Non-Crime Outcomes: Education, Income, and Unemployment using Leave Out Annual Mean as Instrument

	(1) all	(2) at least 1 pre service crime	(3) no pre service crime	(4) father schooling <= 9 years	(5) father schooling > 9 years
	Dep var = schooling > 12 years				
military_service_tax_record	-0.04281	-0.18982**	-0.00094	-0.02337	-0.07493
	[0.03820]	[0.07488]	[0.03705]	[0.04107]	[0.05133]
Mean Dep Variable	0.42	0.22	0.45	0.27	0.47
	Dep Var = any years unemployment benefit 23-34				
military_service_tax_record	-0.03847	-0.08462	-0.04047	0.06615	-0.07569**
	[0.04011]	[0.11262]	[0.03974]	[0.08478]	[0.03625]
Mean Dep Variable	0.46	0.55	0.45	0.5	0.044
	Dep var = any welfare from 23-34				
military_service_tax_record	-0.01569	-0.00148	-0.02922	-0.03618	-0.00478
	[0.02815]	[0.08131]	[0.02805]	[0.04969]	[0.03083]
Mean Dep Variable	0.13	0.25	0.11	0.15	0.12
	Dep var = any_disability pension 23-34				
military_service_tax_record	-0.01007	0.00202	-0.01302*	-0.02966	-0.0017
	[0.00808]	[0.02957]	[0.00739]	[0.01854]	[0.00854]
Mean Dep Variable	0.012	0.024	0.01	0.014	0.011
	Dep Var = # sick days 23-34				
military_service_tax_record	-38.0415**	-9.6126	-45.0625***	6.72722	-54.701***
	[19.07970]	[57.37456]	[16.76052]	[34.34662]	[17.51541]
Mean Dep Variable	48.8	83.7	43.5	59.1	45.2

	(6) all	(7) at least 1 pre service crime	(8) no pre service crime	(9) father schooling <= 9 years	(10) father schooling > 9 years
	Dep Var = income_30_34				
military_service_tax_record	0.13354**	-0.20091	0.19832***	-0.08936	0.21911***
	[0.05750]	[0.21927]	[0.05492]	[0.10451]	[0.06422]
Mean Dep Variable	12.38	12.25	12.4	12.35	12.4
	Dep Var = # years unemployment benefit 23-34				
military_service_tax_record	0.04471	-0.37996	0.07118	0.64869**	-0.17438
	[0.15338]	[0.47138]	[0.15222]	[0.31720]	[0.13764]
Mean Dep Variable	1.55	2.04	1.48	1.79	1.47
	Dep Var = months welfare 23-34				
military_service_tax_record	0.1313	5.09642	-0.87501*	-0.24347	0.36476
	[0.66984]	[3.44656]	[0.49695]	[1.28972]	[0.78766]
Mean Dep Variable	1.56	4.2	1.15	1.98	1.41
	Dep Var = # years with full/partial disability pension				
military_service_tax_record	-0.04169	-0.00419	-0.05293	-0.1677**	0.01267
	[0.03753]	[0.15600]	[0.03964]	[0.07823]	[0.04498]
Mean Dep Variable	0.057	0.113	0.048	0.067	0.053

Robust standard errors, clustered on officiator. *** 1%, ** 5%, * 10*. F-statistic on first stage regressions: 47 for sample with criminal history, 80 for sample without criminal history, 63 for low education fathers and 78 for high education fathers. Each regression controls for test year dummies, county dummies and test year x county dummies, and instruments for military service with the leave out annual mean. Samples sizes are N=168806 for the whole sample, 22590 for criminal history, 146216 for no criminal history, 43904 for father schooling<=9, and 124902 for father schooling > 9 samples for all outcomes other than income. For income, about 3500 observations are missing.

Appendix Table 6. A Comparison of Descriptive Statistics: IV Sample vs. Matched Sample

	IV Sample *N = 168,818*	Matched Sample N = 7,366
Pre-test day controls		
Any weapons < 18?	0.01	0.00
Any violent < 18?	0.02	0.01
Any traffic < 18?	0.06	0.02
Any theft < 18?	0.06	0.04
Any other < 18?	0.04	0.02
Any drugs< 18	0.00	0.00
schooling father	11.14	11.74
schooling mother	11.55	12.26
income father	12.32	12.33
income mother	11.87	11.95
2 year school?	0.18	0.05
3 year school?	0.76	0.92
Test day controls		
Height	179.75	180.12
Weight	71.59	72.51
BMI	22.14	22.25
Ability Score	5.17	5.61
Physical Capacity Score	6.26	6.28