Workplace Drug Testing and Prescription Opioid Use*

Carson Homme
University of Notre Dame

Alec Hoover
University of California San Diego

July 27, 2026

Opioid use disorder cost an estimated $2.4 trillion in United States in 2023 (House 2025). In this paper, we utilize exogenous exposure of US Army soldiers to varying levels of drug testing to explore the potential for workplace drug testing to cause individuals to substitute from illicit drugs to prescription opioids. Being assigned to a high testing environment increases the likelihood that a soldier is drug tested by 12% and decreases the likelihood that a soldier uses illicit drugs by 12%, split between a decrease in marijuana and illicit opioid use. Soldiers exposed to a high testing environment are also more likely to receive an opioid prescription offsetting approximately 78% of the decrease in illicit opioid use. We provide robustness checks that suggest the effect of being assigned to a high testing environment is due to the increase probability of being tested and not other characteristics of the environment - e.g. characteristics of peers or the company commander. This highlights the importance of considering substitution effects when analyzing the effectiveness of drug testing policy.

*All views, opinions, and interpretations are those of the authors and do not represent the views or official positions of the United States Military Academy, the United States Army, or the Department of War.

1 Introduction

Substance use is an increasingly costly issue for Americans. The National Institute on Drug Abuse estimated that illicit drugs cost the US $193 billion dollars in 2007 while prescription opioids cost $78.5 billion in 2013 (Services 2024). The opioid crisis has made this problem dramatically worse. According to the Drug Enforcement Agency, 5.7 million Americans had opioid use disorder and 74,702 Americans died from opioids in 2023. This cost an estimated $2.4 trillion, mostly in lost life and reduced quality of life (House 2025). In addition to the cost to individuals who have died from or continue to live with substance use disorder, substance use has a significant impact on firms and worker productivity. Using data from the National Survey on Drug Use and Health, Ghimire, Florence, and Guy (2025) found that substance use disorder led to an estimated $92.65 billion dollars in productivity related losses in the United States in 2023. This paper examines the effectiveness of one tool used to curb the cost of substance use - workplace drug testing. Specifically, we explore the impact of increased workplace drug testing on both illicit drug use and legal substitutes including prescription opioids.

Using plausibly exogenous variation in exposure to random workplace drug testing for soldiers in the United States Army, we explore the impact of increased drug testing exposure on substance use. Army policy requires companies to randomly drug test a part of the company every month. We find that being assigned to a company commander who averages one additional testing day per month increases the likelihood that a soldier is drug tested by 12% (.067 pp. per day). This also leads to a 12% decrease (.042 pp. per day) in the likelihood of using an illicit drug which is split evenly between a decrease in marijuana use and illicit opioid use. Using data on soldier prescriptions we can also examine the impact on the likelihood of a soldier receiving a prescription. We find limited evidence that soldiers assigned to high testing commanders are more likely to receive any prescription. However, one additional testing day per month increases the likelihood of receiving a new opioid prescription by 2.5% (.0018 pp.) and increases the likelihood of having an active opioid prescription by 3% (.0175 pp. per day). This suggests that approximately 78% of the decrease in illicit opioid use is offset by an increase in prescription opioid use. Interestingly, the increase in opioid prescriptions is driven almost entirely be an increase in oxycodone prescriptions. To add to this narrative of increased drug testing leading soldiers to substitute towards legal substances, we use military police records to examine the impact of increased testing on criminal behavior and alcohol abuse. Here, we again find a positive effect. One additional test per month is associated with a 4%

increase in the likelihood of having a misdemeanor criminal offense. There is no impact on the likelihood of having a felony offense.

While we show evidence that company testing frequency is uncorrelated with individual soldier characteristics, we cannot preclude that drug testing policy is not related to other characteristics of the individual's environment which may independently impact individual substance use. In attempt to isolate the impact of drug testing, we can control for other characteristics of the company and the company commander. Notably we control for measures of commander quality as well as the company's prior drug and alcohol use as well as company lagged prescription opioid use, potential sources of peer spillovers. We see no changes in the estimated impact of increased drug testing frequency. This suggests that are results are not being driven by correlated commander or peer characteristics.

Our hypothesis relies on the idea that illicit drugs have legal substitutes. One struggle with classifying substances as legal or illegal is that similar substances can be found in both categories and some substances can be either legal or illegal depending on how they are accessed. This can create challenges in attempts to deter the use of illegal drugs, as many illegal drugs have very close legal substitutes. Opioids are a prime example of this. Fentanyl and morphine are legally used in medical settings, but are illegal when possessed without medical documentation. The effect of drug testing to deter substance use is limited by the extent to which individuals substitute from illegal substances to legal substances. Drug substitution has been documented before, particularly among marijuana users. Analysis of both survey data and state legalization laws in the US has found marijuana to be a substitute for alcohol, opioids, anti-depressants, and anti-anxiety medication (Carrieri, Madio, and Principe 2020, Lucas and Walsh 2017, Lucas et al. 2016, and Perrone, Helgesen, and Fischer 2013). Beyond marijuana, there is also some evidence that prescription opioids and illicit opioids are substitutes while illicit opioids and cocaine are not (Alpert, Powell, and Liccardo Pacula 2018). Our analysis includes soldier prescription opioid use and alcohol abuse. While it would be interesting to look at evidence of substitutes for cocaine or amphetamines, we do not have data on common stimulants.[1]

Our paper will also provide new evidence on the responsiveness of drug users to moderate changes in workplace drug testing policies. Previous research into workplace drug testing has found that drug testing is correlated with decreased self-reported mari-

1. Individuals within our setting do not receive prescriptions for stimulants including Adderall or Ritalin. We cannot observe individual caffeine or nicotine consumption

juana use and prescription pain reliever misuse (Carpenter 2006 and Lee and Ross 2011); however, prior studies have lacked exogenous variation in drug testing policies as well as administrative data on substance use.

The rest of the paper is as follows: section 2 discusses the institutional details of the Army and Army drug testing policy, section 3 describes the data and sample, section 4 introduces a model of drug use, section 5 covers the empirical strategy, section 6 shows the results, section 7 examines mechanisms, section 8 provides some robustness checks and section 9 concludes.

2 Institutional Details

2.1 Drug Testing in the Army

The Department of Defense prohibits drug use for all employees (Army Regulation 600-85). Article 112a of the *Uniform Code of Military Justice* prohibits soldiers from using amphetamines, barbiturates, cocaine, ecstasy, opiates, heroin, phencyclidine (PCP), tetrahydrocannabinol (THC), oxycodone/oxymorphone, benzodiazepines, lysergic acid diethylamide (LSD), steroids, synthetic cannabis (Spice), and any related compounds. Before entering service, all soldiers must pass a drug test and remain subject to ongoing urine drug testing throughout their military service. Failing a drug test carries significant career consequences. Army Regulation specifically mandates that commanders must initiate separation for every soldier that fails a drug test (Army Unit Prevention Leader's Handbook). This process is costly, both to the Army and commanders, with drug testing failure-related separations accounting for 19% of all soldiers who are separated from the Army for misconduct.[2]

2.1.1 Company Commanders in Army Drug Testing

Battalions are the smallest echelon of Army units responsible for instituting a drug testing program (Army Regulation 600-85). Yet, Company Commanders are the lowest-level leaders responsible for implementing and enforcing this program. Company Commanders decide when and how many soldiers are tested to meet program requirements. Battalion Commanders oversee the company's implementation and can administer larger

2. Failed drug tests account for 73% of all drug crimes within the Army.

punishment (including separation) for drug offenses. Battalion Commanders are the first level of approval to separate soldiers on drug charges and adjudicate punishment to soldiers for these crimes under the Uniformed Code of Military Justice. Company commanders (the next echelon of unit below) are responsible for scheduling and executing the drug testing within their units.[3] Our study focuses on the effect of various Company Commanders' testing policies on drug use.

2.1.2 Types of Tests

Company Commanders (henceforth "commanders") must test at least 10% of their personnel monthly, and each soldier must be tested at least once a year (Army Unit Prevention Leader's Handbook). These tests take the form of commanders either testing a random fraction of service members in their unit (an "Inspection Random" or "IR" test) or commanders testing all soldiers in their formation (an "Inspection Unit" or "IU" test). Commanders cannot directly select individual soldiers to test without probable cause justified by an Army lawyer. While soldiers do not have a choice over whether to participate or not, commanders may excuse a soldier from testing. Reasons for missing a drug tests can vary, such as soldiers being on vacation, being tasked with additional duties outside of their unit footprint, or being within an occupation with different work hours than the rest of the company. If a soldier cannot test when selected, Army regulation mandates that commanders test this soldier at his or her next availability using an "Inspection Other" or an "IO" test code (Department of the Army 2016).[4] For example, cooks within the Army often work longer and very different hours due to the need to provide breakfast, lunch, and dinner. Their ability to get away from their job to participate in unit drug testing is constrained, and the commander may excuse them on the test date and reschedule for later if they are selected for testing.

Selection for "IU" tests includes all soldiers within the company, including the commander, and all tests are completed on a single day. Selection for "IR" tests occurs through a computerized randomization process where all soldiers have an equal probability of being chosen and also occurs on a single day.[5] Commanders can use either type of test to

3. Figure A1 within the appendix helps illustrate this hierarchy within Army units.

4. Any soldier not tested within the first three quarters of the fiscal year is required to be tested under this code at some point within the fourth quarter (Army Unit Prevention Leader's Handbook).

5. If the computer program (DTP Lite software provided by the Army Substance Abuse Program) is unavailable, Army policy suggests alternative randomization methods, such as drawing index cards from a hat or using a 10-sided die.

meet the monthly 10% requirement. It is essential to note that each "IR" test is akin to a draw with replacement. A soldier getting tested under "IR" conditions on the first of the month has no effect on the probability of that soldier getting "IR" tested at another drug test later in the month.

2.1.3 Drug Testing and Prescription Drugs

Army drug testing policy includes an exception for soldiers who test positive with medical documentation (Department of the Army 2016). Upon testing positive soldiers can present evidence of proper medical use, usually a doctor's note or a prescription, of the substance for which they tested positive. This is particularly important for opioids. The only opiate that the Army specifically tests for is heroin. If a soldier tests positive for an opiate, not heroin, and they present a prescription or doctor's note for an opioid, the test will be recorded as a negative. Unlike in other settings during this time period, the Army operates under federal drug law. This means that soldiers cannot receive medical marijuana - that is soldiers cannot present evidence of proper medical use of marijuana.

2.2 How Soldiers and Commanders Are Assigned to Companies

The Army's structured assignment of soldiers and commanders minimizes self-selection biases, such as drug-using soldiers choosing commanders who avoid testing, as neither can easily influence their placement. The Army's Human Resources Command assigns soldiers to brigades, which consist of multiple battalions, based primarily on U.S. Army personnel requirements.[6] Once a soldier arrives at a brigade, the brigade's HR office (also known as the Brigade "Adjutant's Office" or the Brigade "S-1 Shop") is responsible for establishing a process that assigns each soldier to a specific battalion and company according to the soldier's occupation and rank in a manner that generally keeps all subordinate units at similar personnel strength (Army Regulation 614-200). In practice, battalion and company personnel strengths fluctuate daily and are often imperfect due to lags in the time it takes for personnel managers to update soldiers' records. As a general rule, soldiers have little ability to influence the precise company they are assigned to (Lyle 2006). Importantly, it is highly unlikely that the soldiers in our sample would have prior knowledge of battalion or company drug testing procedures, much less the ability to sort into

6. See Section 2.1 of (Bruhn et al. 2024) and (Army Regulation 614-200) for additional details on how the Army assigns soldiers to brigades.

units based on such information.

Company commanders are assigned to brigades through the above process; however, when an officer is assigned command of a company depends on their commissioning year group, occupation, and arrival at the brigade. For most occupations, company command is a required job officers must complete to get promoted to the next rank. Officers generally conduct this job as captains with five years of service and are considered for promotion afterward when their year group reaches nine years of service. The Army implements an "up-or-out" promotion system and evaluates all officers for promotion at certain years of service relative to the year they were commissioned. This implies that officers must complete company command before their eighth year of service to be competitive for promotion; thus, outside of occupation, officers are assigned commands primarily based on their timeline before promotion, and officers must serve in command for at least 12 months to be competitive for promotion. Thus, we argue that commanders within our sample are assigned orthogonally to a company's drug use, officer quality, or other measure of "fit."[7]

3 Data

The data for this project come from the Office of Economic and Manpower Analysis (OEMA) within the United States Military Academy at West Point. The sources include administrative, evaluation, drug testing, and crime. We use the administrative data for a soldier's demography, education, and occupation. The evaluation data contain annual performance reviews for every officer in service. We use this data to identify Company Commanders within the sample. The monthly pay data comes from the Defense Financial Accounting Services and lists a soldier's pay, assignment unit, deployment status, and rank. The crime data comes from the Army Law Enforcement Reporting Tracking System (ALERTS) database that military law enforcement personnel oversee. Critical to note is that this database contains criminal incidents that occur both on and off a military installation.[8] Finally, the drug testing data comes from the Drug and Alcohol Manage-

7. We will conduct a robustness test to see if measures of officer ability/quality are correlated with company drug testing or drug use.

8. Off-post police forces report crimes external to a military post directly to the military police. The military police then input these crimes into the ALERTS database. Soldiers can commit crimes and they never get reported into ALERTS; thus, for all off-post crimes reported, we interpret these events as a lower bound for crimes committed off-post.

ment Information System (DAMIS) that the Army Substance Abuse Program manages. Of note, the ALERTS and DAMIS databases do not feed into one another. Suppose a soldier tests positive for drug use and is logged within the DAMIS system after the sample is processed. In that case, logging that soldier's criminal offense within the ALERTS database will take a separate and additional action. Likewise, if law enforcement (either military or civilian) catches a soldier with drugs and then the Army administers that soldier a drug test, that soldier is first logged into the ALERTS database and his or her resultant drug test is logged within the DAMIS database. Noteworthy within the context of this study is that these separate sources imply the treatment (a soldier's drug test) and crime outcomes (a criminal offense incident not related to testing) come from different data sources.

3.1 Sample Construction

We draw our sample from all enlisted soldiers where we can identify the commanders from January 2015 through April 2022. To ensure we have enough time and consistency to identify how a commander's testing impacts drug use, we make the following restrictions. First, we include only company commands where the commander serves for 12 consecutive months without deploying and we drop all soldiers who are deployed. Drug testing and drug use are predominantly tasks that are conducted in a non-deployed environment. Moreover, a deployed environment distorts regular incentives and is not externally valid to other workplace drug testing programs.

Second, we restrict our sample to only "operational units," which are those units that will actually deploy. "Non-operational" or "training" units are those like basic training or recruiting companies that are responsible for generating and training the soldiers that will fill the operational force. Soldiers often do not serve long within these units, and the composition and authority structure is unlike the traditional Army structure outlined in Figure A1.

Third, we restrict our sample to commands that average at least 40 and less than 225 soldiers. During this time period, the Army was shutting down certain units and setting up new ones; thus, this sample restriction helps us eliminate transitional units that could have potentially held a selected sample of injured soldiers or soldiers facing other legal or personal issues that would prevent them from being moved.

Last, we drop soldiers after their first drug or alcohol offense. Since a drug or alcohol offense is career-ending for most soldiers, getting caught once reduces the cost of more punishment for a soldier, thereby increasing the probability of using again in the future. To accurately identify the effect of a commander's testing on a soldier's drug use, we need to be able to compare soldiers of similar incentive structures and capture an outcome that accurately measures the underlying outcome.[9]

After implementing the above restrictions, we are left with a sample of 323,777 soldiers and 7,488 commanders. Table A2 shows summary statistics. Our sample contains mostly young men. Soldiers have an average of 2.15 drug tests per year with less than one percent of soldiers having a positive drug test or drug related criminal offense. Commanders in our sample have served an average of 7.29 years at the start of their command and stay in command for just under 16 months.

4 Modeling Substance Use

The decision over when and what drugs to use can be thought of under the framework of a Becker Crime Model (Becker 1968). Individuals choose their total drug consumption in each period to maximize their individual specific utility, $U_i(Drug_{i,t})$. The utility of using a bundle of drugs is the combination of the direct benefits received from the drug use and the expected cost of the drug use. The expected cost to testing positive for a banned substance is being discharged from the Army.

$$U_i(Drug_{i,t}) = u_i(Drug_{i,t}) - \mathbb{P}[TestPositive_{i,t}] * c_i(discharge_{i,t}) \tag{1}$$

The probability of getting caught using drugs is the probability of being tested on day t and testing positive on day t given the individual's prior drug use.

$$\mathbb{P}[TestPositive_{i,t}] = \mathbb{P}[tested_{i,t} \cap Positive_t | Drug_{i,t}, Drug_{i,t-1}, \dots] \tag{2}$$

If a random sample of soldiers are selected to be tested, then

$$\mathbb{P}[tested_{i,t} \cap Positive_t | Drug_{i,t}, Drug_{i,t-1}, \dots] = \mathbb{P}[tested_{i,t}] * \mathbb{P}[Positive_t | Drug_{i,t}, Drug_{i,t-1}, \dots] \tag{3}$$

We identify exogenous variation in $\mathbb{P}[tested_{i,t}]$. This variation will impact the deterrence cost of drug use if that drug use impacts the likelihood of the soldier testing positive.

9. This restriction introduces a negative mechanical bias into our estimates of the impact of increased testing on drug use. The size of this bias is discussed in the Appendix Section B.2.

This leads to two related observations. First, an increase in testing does not decrease the utility of using prescription drugs. Second, an increase in testing will decrease the utility of using banned substances in proportion to the length of time that substance remains detectable in a drug test. That is, an increase in testing should have a larger impact on the use of marijuana, which is detectable for 7-21 days after last use, than on Heroin, which is detectable for 2-3 days after last use (Moeller, Lee, and Kissack 2008). Assume that individuals have the choice of using three substances: Marijuana, Fentanyl, and Oxycontin - $Drug_{it} = \{W, F, O\}$.[10] Here an increase in testing leads to the greatest increase in the

Drug Use	Probability of Testing Positive (Deterrence Cost)
$Drug_{i,t} = \{W = 1, F, O\}$	Probability of being tested in the next 7-21 days
$Drug_{i,t} = \{W = 0, F = 1, O = 0\}$	Probability of being tested in the next 2-3 days
$Drug_{i,t} = \{W = 0, F, O = 1\}$	0
$Drug_{i,t} = \{W = 0, F = 0, O\}$	0

probability of testing positive in the future if the individual uses marijuana regardless of other substance use. It has the next greatest impact on individuals who use Fentanyl but do not use Oxycontin. Testing has no impact on the likelihood of testing positive for individuals who do not use marijuana and either do not use Fentanyl or use Fentanyl and Oxycontin as these individuals would not test positive even if they were tested. Though this suggests that the impact of increased testing could be greater for marijuana than opioids or other drugs that are not detectable for as long, the actual change in substance use depends on the total utility from the substance use as well as how close of a substitute there is for the preferred drug. The only substance with a direct prescription substitute in our setting will be opioids as soldiers are not allowed to use medical marijuana and no one in our sample receives a prescription for amphetamines such as Adderall or Ritalin.

5 Empirical Strategy

Our goal is to identify exogenous variation in the probability that a soldier is tested. To do this we rely on the conditional random assignment of soldiers to commanders with different propensities for conducting drug tests.

10. Oxycontin represents a prescription opioid. The drug test cannot distinguish between Oxycontin and Fentanyl, a type of opioid. Therefore, if an individual has a prescription for Oxycontin, they cannot test positive for Fentanyl.

5.1 Estimating Commander Testing Frequency

To identify each commander's proclivity for drug testing, we calculate the number of days a commander conducts an inspection drug test each month. [11]

$$CDR_j = \overline{NumTestingDay_{c(i),t}}$$

In practice we estimate the commander fixed effect, CDR_j, using the first four months of the commander's tenure. In order to improve the relevance of the measure, we then update it based on the month of the command that each soldier arrives at the company using the 4 months prior soldier's arrival. For example a soldier that arrives during the 10^{th} month of the command will be assigned a value of the fixed effect for their commander estimated using months 6-9 of the command, $CDR_{j(i)} = CDR_{j,10}$. Meanwhile a soldier that arrives in month 6 will have a value $CDR_{j(i)} = CDR_{j,6}$.[12] Figure 1 shows the distribution of commander testing frequency. The plurality of commanders average 1 inspection test per month with few commanders averaging more than two tests per month.

5.2 Estimating Impact of Testing Frequency on Soldier Outcomes

We estimate the impact of commander testing frequency on soldier outcomes using the following equation.

$$Y_{i,t} = \alpha + \beta CDR_{j(i)} + V_i + \omega_{o,g,l,q} + \zeta_m + \varepsilon_{i,t}$$

Here $Y_{i,t}$ are the soldier outcomes on day t. This includes being tested, testing positive, receiving a prescription, or having a criminal offense. We include fixed effects, $\omega_{o,l,g,q}$ for the interaction of soldier occupation, grade, location, and quarter of arrival at the company as well as the calendar year-month, ζ_m. We will also include in controls for individual demographics V_i. Identification primarily relies on the commander testing frequency, $CDR_{j(i)}$ being uncorrelated with a soldier's propensity to use different substances.

11. Alternatively, we have also estimated the commander fixed effect using the percent of a company tested each month (testing volume) instead of number of tests the company had in the month (testing frequency). Testing frequency is a better predictor of the likelihood that soldiers are tested in the future

12. We have calculated commander testing frequency using different periods of time (e.g. the average over 6 months instead of 4 months). Figure 4 panel B shows that the results are similar. Using a longer period of time to calculate commander testing frequency reduces the number of soldiers in our estimation sample.

5.2.1 Covariate Balance

Formally the identifying assumption is

$$cov(CDR_{j(i)}, \varepsilon_{i,t} | \omega_{o,g,l,q}, \zeta_m) = 0$$

This assumption rests on the idea that soldiers are conditionally randomly assigned to companies with commanders of varying testing propensities. While this guarantees that soldier characteristics are independent of company and company commander characteristics, we only need the weaker assumption of zero conditional covariance. Based on Army regulation, HRC can consider occupation, grade group (E01-E04, E05-E07, and E08-E09), contract length, and soldier location preferences when assigning soldiers. Therefore, conditional on these characteristics soldier characteristics should be uncorrelated with commander testing frequency.[13]

We test this assumption by regressing soldier characteristics on the commander testing frequency.

$$V_i = \alpha + \beta CDR_{j(i)} + \omega_{o,g,l,q} + \zeta_m + \theta_{i,t} \tag{4}$$

Table 1, panel A shows the results. The table suggests that commander's that conduct more frequent drug testing are more likely to be assigned black soldiers.[14] This correlation is relatively small. Commanders that do 1 additional test per month have an average of 0.5 additional black soldiers in a company of 100-150 soldiers. Additionally, commander testing frequency is not correlated with soldiers having a moral waiver. Individuals with a history of drug use or criminal activity must obtain a waiver to enlist in the Army. All soldiers must pass a drug test when they enlist. So having a waiver is the only observable measure of risk of soldier drug use. This provides some evidence that soldiers are not assigned to commanders based on the Army's perception of the likelihood that the soldier will use drugs. We will show the impact of controlling for these soldier characteristics in our results to address concerns with non random assignment of soldiers to commanders.

13. Other minor considerations include joint military couples and EFMP which primarily affect the location a soldier can be assigned. This portion of our sample is small, and we do not control for it within our specifications outside of controlling for soldier location.

14. We have been unable to explain this correlation. Controlling for commander characteristics including race and gender as well as restricting to junior enlisted soldiers or more standardized units - brigade combat teams - does not make a difference.

5.2.2 Correlated Commander Characteristics

Random assignment of soldiers to commanders ensures that our reduced form estimates of the impact of being assigned to a high testing commander are unbiased. However, if we want to interpret our results as being the impact of being exposed to increased drug testing, we need an assumption similar to an exclusion restriction. That is, we need to show that our measure of commander testing frequency is not correlated with other commander or company characteristics that may influence a soldier's likelihood of using drugs. We have two primary types of concerns. The first is that high testing commanders or companies with high testing commanders have some other characteristic in common that decreases illicit drug use separate from drug testing. Our other concern is that high testing commanders engage in some behavior that encourages soldiers to seek prescription drugs or other legal substitutes for illicit drugs. This would create a substitution effect that is not due to increased drug testing.

Table 2 Panel A starts by showing the correlation between our measure of commander testing frequency and commander characteristics. High testing commanders - commanders with higher values of commander testing frequency - are less likely to be Hispanic. Commander testing frequency is uncorrelated with other commander demographics including, gender, marital status, education, age, or years in the Army. One might be concerned if commander testing frequency is correlated with commander quality. It is possible that high quality commanders are better at encouraging their soldiers to follow the rules - not use illicit drugs - or prioritize their soldier's health which could increase soldier medical utilization and prescription opioid use while decreasing illicit drug use. Table 2 Panel B shows the relationship between commander testing frequency and some measures of officer quality. Testing frequency is not correlated with measures of the commander's educational quality or ability; however, high testing commanders receive on average a higher percentage of good performance ratings and are more likely to be promoted to Major early.[15] This suggests that our reduced form estimates may not be solely due an increase in drug testing, but also the effect of having a higher quality commander. A related concern is that commander testing frequency is correlated with characteristics of the company. Suppose having a larger percentage of a company get caught using illicit drugs or having a prescription of opioids causes a commander to do more drug testing to try to deter future illicit drug use. If there are any peer spillovers from having more

15. Army performance reviews rate service members on a scale of 1-4. The percent of reviews in which a service member receives a 4 or "Most Qualified" is a common statistic used for officer promotions.

peers within the company using drugs, getting caught using drugs, or having a prescription opioid, the relationship between commander testing frequency and soldier outcomes would reflect both the impact of increased testing and the potential peer spillovers. One way we can start to get at this is to construct measures of company drug use, prescription use, and health status, similar to our measure of commander test frequency. We construct the average number of soldiers in commander j's company in each month with any prescription, opioid prescription, positive drug test, or health rating and then take the average over four months.[16] We define $\overline{X}_{j,m} = \frac{1}{4}\sum_{\tau=m-4}^{m-1} \frac{\sum_{i \in c,\tau} X_{i,\tau}}{N_{c,\tau}}$ where $N_{c,\tau}$ equals the number of soldiers in company c during month τ. To reduce potential concerns with simultaneity, we use the four months prior to the start of the command. We run the following regression.

$$\bar{X}_{j,1} = \delta_0 + \delta_1 CDR_{j,m} + \mu_{j,m} \tag{5}$$

Table 2 Panel C shows that high testing commanders are more likely to be assigned to units that previously had more frequent drug tests, more positive drug tests, more alcohol criminal offenses and more prescriptions. This further cautions us against interpreting the reduced form estimates as solely due to an increase in the likelihood of being drug tested as commander testing frequency may be correlated with important peer spillovers.[17] Section 7 will attempt to separate out the channels through which being assigned to a high testing commander may impact soldier substance use.

5.2.3 Estimating the impact of testing on drug use

When attempting to estimate the impact of changes in the probability of being tested on the likelihood that a soldier uses a banned substance, we face the problem of not observing all soldiers drug use. Instead, we only see when a soldier tests positive for a banned substance. To recover the impact of testing frequency on the use of banned substances we need to make an additional assumption on how soldiers are selected for drug testing. That is, we need to assume that being selected for drug testing is independent of

16. Soldiers health is related along six dimensions of physical and mental health to determine their ability to carryout certain duties.

17. Note that the company characteristics under the prior commander are not a perfect peer measure. When commanders start their command, they are instructed to not make any immediate major changes. Therefore, they may inherit some of the testing practices of the previous command, particularly at the beginning of their command. An increase in the prior commander's testing may lead to both a change in company drug use and the current commander's testing frequency.

whether the soldier uses drugs.

$$Test_{i,t} \perp\!\!\!\perp DrugUse_i$$

Being selected for an inspection test should be independent of soldier drug use. As discussed previously, inspection tests consist of inspection random tests which are conducted as a random sample, inspection unit tests in which all soldiers in the unit are tested, and inspection other tests which primarily consist of testing soldiers selected for inspection random and unit tests who were unable to be tested at the selected time. We provide evidence in support of this claim in table 1, panel B. We regress each soldier characteristic on a dummy variable for having an inspection test on day t.

$$V_{i,t} = test_{i,t} + \omega_{o,g,l,q} + \zeta_m + \vartheta_{i,t}$$

Clearly, inspection tests are not random. They are, however, not random along predictable lines. Female soldiers and soldiers with dependents (married or have children) are less likely to be tested. This is due to requirements on having a same-gender observer for tests as well as soldiers with dependents being less likely to live on base and being available to test when not on duty. Panel C adds controls for gender and having a dependent. While testing is still correlated with race, the p-value that the coefficients is jointly significant is 0.12. Additionally, the fact that being tested is not correlated with the soldier enlisting with a moral wavier suggests that inspection tests do not target soldiers with a higher risk of drug use.[18]

The reason that we need tests that are independent of drug use is that an increase in the likelihood of being tested increases the likelihood of testing positive mechanically separate from any potential change in substance use. However, since Inspection tests are independent of substance use we can separate the mechanical effect.

$$\mathbb{E}[DrugUse_i] = \mathbb{E}[DrugUse_i | tested_{i,t} = 1] \tag{6}$$

$$= \mathbb{E}[tested_{i,t} \cap DrugUse_i | tested_{i,t} = 1]$$

$$= \mathbb{E}[TestPositive_{i,t} | tested_{i,t} = 1]$$

18. Table A3 repeats table 1 panel C only using Inspection Random and Inspection Unit tests. There may be some concern that commanders can selectively enforce Inspection Other tests. Additionally, it is possible that soldiers who miss an IR/IU test may be able to alter their behavior in between learning of the missed tested and being tested (IO) themselves. The results are similar, however we do get a larger p-value.

Thus, to estimate the impact of testing on the use of banned substances we can restrict our sample to soldiers who had an inspection test.[19]

6 Results

6.1 Impact on Illegal Substance Use

The main question of this paper is whether the deterrence effect of drug testing is crowded out by the potential substitution effects. The first step in answering this is to measure the impact of increased drug testing on illicit drug use. Figure 2 panels (a) and (b) shows the raw correlation between the estimated commander testing frequency and the likelihood that a solider is tested and tests positive. Panel (a) shows that there is a relatively strong positive linear relationship between commander testing frequency and the likelihood of being tested; however, Panel (b) shows that there is only a weak positive relationship between commander testing frequency and testing positive. Note that the correlation in panel (b) includes both the mechanical effect of increased testing and the potential behavior response to the increased probability of being tested. Panel (c) shows the relationship between commander testing frequency and the likelihood of testing positive conditional on having an inspection test. As previously discussed the likelihood of testing positive conditional on being tested is in expectations equal to the likelihood that a soldier uses drugs. Panel (c) shows a negative correlation between commander testing frequency and any drug use. Panels (d)-(f) show that commander testing frequency is negatively correlated with the three most common drugs that soldiers test positive for, marijuana, cocaine, and opioids.[20]

Table 3 shows the results of the main specification, equation 5.2. Panel A shows the impact of a 1 unit increase in commander testing frequency on the likelihood that the soldier is tested while Panel B shows the impact on drug use. Column (3) is our primary specification with brigade included in the fixed effects and controls for individual covariates. One additional testing day per month leads to a .0729 percentage point (or 12.8%) increase in the likelihood that a soldier is tested on a given day. Meanwhile, the same

19. The Math Appendix shows that we can equivalently recover the impact of testing frequency on drug use by controlling for whether a soldier is tested and then dividing the coefficient by the probability of being tested.

20. Figure A2 shows the relationship between commander testing frequency and drug use after controlling for the year-month and the interaction of soldier brigade, rank, occupation, and quarter of arrival at the company - the fixed effects in our main specification.

increase in commander testing frequency leads to a .0428 pp. (11.9%) decrease in the likelihood that a soldier uses drugs.[21] We prefer using Brigade in our fixed effects instead of location as in column (1) as drug testing is likely correlated with a unit's daily activities. Units with more time to conduct drug testing may also have more time to attend medical appoints and receive prescriptions. Comparing columns (1) and (2) shows that controlling for brigade does reduce impact of commander testing frequency on the likelihood of being tested. This makes sense as drug testing policy is set in part at the brigade level. Comparing columns (2) and (3) shows that controlling for individual characteristics does not impact either estimate. This should help resolve any lingering concerns that soldiers are systematically assigned to high or low testing commanders.

Table 4 shows the impact of commander testing on drug use by type of drug. A 1 unit increase in commander testing frequency decreases opioid use by a statistically significant 0.0223 pp. (43.7%). The impact on marijuana is a similar size, -0.0219 pp. (11.6%), but statistically insignificant while the impact on cocaine use is small and statistically insignificant.[22] Based on our model we would expect the decrease in use of each type of drug to be relative to the length of time the drug remains in an individual's system after use as well as the availability of legal substitutes. Marijuana is detectable for longer than cocaine or opioids so we would expect the decrease in use to be greater for marijuana. However, relative to the respective means, the is the largest decrease in opioid use. This could be due to illicit and prescription opioids being very close substitutes.

6.2 Impact on Substitute Behaviors

While the previous section has shown that an increase in workplace drug testing decreases the use of illicit drugs, the impact of drug testing depends on the extent to which individuals change their behavior to use similar substances. We are going to measure these substitution effects in two ways. First, we will use soldier electronic health records to look at the impact of increased testing frequency on prescriptions with a focus on

21. Note that the sampling with replacement nature of drug testing introduces a negative mechanical bias into this estimate. Commanders that do more drug testing at the start of their command will mechanically catch and remove a larger percentage of the drug users in their companies. Future drugs tests will now have a small percentage of positive tests even without any behavioral response. The Math Appendix provides attempts to calculate the mechanical bias. The estimated impact of commander testing frequency on drug use appears to be 3-10% to large (table B1).

22. Table A4 repeats tale 4 restricting to IR and IU tests. We get slightly larger estimates for the impact of commander testing frequency on drug use. This addresses potential concerns that soldiers may have prior knowledge of IO tests or that commanders may be able to influence IO tests.

opioid prescriptions. The previous section confirmed our model's prediction that drug testing should have the largest impact on the illicit use of marijuana and opioids. Prior research would suggest that this decrease in marijuana and illicit opioid use could lead to an increase in prescription opioid use (Alpert, Powell, and Liccardo Pacula 2018, Carrieri, Madio, and Principe 2020, Lucas and Walsh 2017, and Lucas et al. 2016). We will supplement this analysis by looking at soldier criminal offenses related to alcohol using Army Criminal Investigation Division records. While we cannot see all alcohol consumption, alcohol related criminal offenses should serve has a measure of extreme or problematic alcohol use.

6.2.1 Impact on Prescription Opioids

Figure 3 shows the relationship between commander testing frequency and having a prescription. Panels A and B show that soldiers with higher testing commanders are more likely to receive any prescription as well as an opioid prescription. While this is initial evidence that increased drug testing may increase prescription opioid use, there may be concerns about confounders biasing the relationship in the raw scatterplot.

Table 5 shows the results from equation 5.2 where the outcome is a binary variable for a soldier receiving a new prescription on day t. Panel A shows the impact of an increase in testing frequency on any new prescription. While there is a positive effect in column (1), when we include brigade in the fixed effects instead of location, this relationship disappears. A potential explanation for this is that drug testing may be more likely to occur at times when the company is less busy. Soldiers may also be more likely to visit the doctor when the company is less busy. When comparing between different brigades in the same location we may be comparing soldiers in companies with very different daily schedules. When only comparing between companies in the same brigade, we reduce the correlation between drug testing and prescriptions that may be due to the company's schedule. Panel B shows the impact of an increase in testing frequency on new opioid prescriptions. As when looking at all prescriptions, controlling for brigade instead of location reduces the point estimate. However, the impact on opioid prescriptions remains statistically significant. Being assigned to a commander that averages 1 more inspection testing day per month increases the likelihood of receiving an opioid prescription on a given day by 0.0022pp (2.8% of the mean).[23]

23. The interpretation of this impact may depend on the type of opioid soldiers receive. Table A5 repeats our main specification, column (3), showing the impact of an increase in commander testing on the

Table 4, found that an increase in testing frequency has a much larger impact on the likelihood of using drugs, decreasing total drug use by 0.042pp and illicit opioid use by 0.022pp. In order to test positive for an illicit opioid, a solider would have had to use the drug within 1-3 days of the drug test. To get a comparable measure of the impact on soldiers using a prescription opioid we look at the impact of testing frequency on the likelihood of a soldier having an active prescription. We determine a soldier to have an active prescription based on the date the prescription was written and the daily supply. For example, if a soldier receives a 7-day prescription on January 1st, they have an active prescription for January 1st-7th. Table A6 shows the impact of increased testing frequency on having an active opioid prescription. Being assigned to a commander who averages 1 additional day of testing per month increases the likelihood of having an active opioid prescription by 0.0175pp (3% of the mean). This suggests that about 78% of the decrease in illicit opioid use is offset by the increase in prescription opioid use. That is not to say that every soldier who was induced to use a prescription opioid because of increased drug testing would have otherwise used illicit opioids. It is possible that the increased in drug testing is pushing some illicit opioid users to not use any illicit or prescription opioid as well as incentivizing some marijuana users to switch to prescription opioids.

6.2.2 Impact on Criminal Offenses

Another way to measure substance use is to use criminal offense records. Unfortunately most drug crimes are tied to a urine test and we cannot identify drug offenses tied to a police drug test separate from offenses tied to a company drug test. This makes it difficult to separate out the impact of company drug testing on criminal drug offenses that is not driven by the mechanical impact of increased testing on increased positive tests. We can however, look at the impact of increased drug testing on alcohol related criminal offenses. Though not a comprehensive measure of alcohol use, an increase in alcohol related criminal offenses would indicate that increased drug testing increases heavy or

likelihood of receiving a prescription for different types of opioids. Note that this is not an exhaustive or mutually exclusive list as soldiers can receive multiple prescriptions at the same time. The most commonly proscribed opioid is oxycodone, followed by hydrocodone and codeine. The impact of increased commander testing frequency on having a new opioid prescription is almost entirely concentrated in new oxycodone prescriptions. Being assigned to a commander that averages 1 more inspection testing day per month increases the likelihood of receiving an oxycodone prescription on a given day by 0.00187pp compared to 0.002pp for any opioid. There is also a small increase in the likelihood of receiving a morphine prescription.

problematic alcohol use.[24]

Table A7 shows the impact of commander testing frequency on the likelihood of having a criminal offense (equation 5.2). Panel A looks at the likelihood of any criminal offense while panels B and C look at misdemeanors and non-violent felonies respectively. Being assigned to a commander that averages 1 additional inspection test day per month increases the likelihood of a soldier having a misdemeanor criminal offense by 0.0006-0.0007pp. (2% of the mean). There is no impact on non-violent felonies. Misdemeanors are important as alcohol offenses make up 38% of misdemeanors. This is suggestive evidence that increased drug testing causes soldiers to substitute not only towards prescription opioids, but also alcohol.

We have shown that increased drug testing increases the use of prescription opioids and criminal offenses. The increase in prescription opioid use is particularly concerning as it offsets most of the decrease in illicit opioid use. For a more complete picture, it would be useful to also look at substitution towards legal stimulants as substitutes for cocaine and amphetamines; however we have no measure of widely used stimulants including caffeine and nicotine and our setting does not include any prescriptions for commonly used prescription stimulants such as Adderall or Ritalin.

7 Mechanism

In the previous section, we showed that being assigned to a commander who has conducted more frequent drug tests increases the likelihood of the soldier being tested in the future, decreases the likelihood of the soldier using illicit drugs and increases the likelihood of the soldier using a prescription opioid and having a misdemeanor criminal offense. We have ascribed this affect to the increase in the likelihood of being drug tested incentivizing soldiers to switch from illicit substances towards legal substitutes. This is, however, not the only possible story. In section 5 we showed that commander testing frequency is correlated with commander quality and the company's prior drug and alcohol use. In the following sections we will attempt to rule out the possibility of the results being driven by correlated commander characteristics as well as peer spillovers.

24. Lucas et al. (2016) found that alcohol is a substitute for marijuana.

7.1 Correlated Commander Behaviors

Even though our narrative is relatively straight-forward, there is a potential concern that the estimated commander testing frequency is correlated with other characteristics of the commander that could influence soldier substance use. This would not invalidate our identification, but it would challenge the interpretation of the results as being the effect of an increase in company drug testing. For example, it is possible the commanders that do a lot of drug testing are also strict in other aspects of their command including physical fitness. These strict commanders could have more soldiers become injured which could increase their use of prescription drugs, particularly opioids. The increased access to prescription drugs could decrease the use of alternative methods of managing pain including marijuana and illegal opioid use. While no test can definitively disprove this type of story, we can provide some evidence against it. We add controls for commander characteristics to regression 5.2. Table A8 columns (2) and (3) shows that including controls for commander demographics and commander quality from table 2 has little impact on the results. This suggests that the results are not being driven by other commander behaviors.

7.2 Peer Spillovers

One may also be concerned with peer spillovers. It is possible that the amount of testing a commander does is responsive to the number of soldiers in the company that are caught using drugs or engaging in other problematic behaviors. If this is the case, then being assigned to a higher testing commander would be correlated with having more peers who use illicit drugs. Typically, if peer spillovers exist, we would expect them to be positive. That is, having more peers that use illicit drugs would increase the likelihood of a soldier using illicit drugs. However, if commander testing frequency is correlated with the number of peers who get caught using drugs, the spillover could be negative as soldiers see more peers facing the consequences of being caught. To map this back into our model of drug use, having more peers caught using drugs could increase a soldier's belief over the likelihood of testing positive conditional on using drugs. This would have the same predicted impact on soldier behaviors as an increase in the probability of a soldier being drug tested.

Addressing this concern is complicated by the fact that an increase testing could create peer spillovers. Under the assumption of positive peer spillovers of substance use, an increase in testing would decrease illicit drug use which would then create a peer spillover that would further decrease illicit substance use. Any attempt to control for the substance use of the company during the commander's command would suffer from issues of simultaneity as testing policy impacts substance use and substance use may impact testing policy. While imperfect, we attempt to control for peer spillovers by adding controls to equation 5.2 for the number of soldiers in the company who had a positive drug test or a criminal offense for drugs or alcohol in the four months prior to the start of the commander's command. This should control for the potential correlation between commander testing frequency and peer substance use. Table A8 shows that adding these controls does not meaningfully impact the estimates for the impact of commander testing frequency on the likelihood of being tested, drug use, or having a new opioid prescription. The estimate for having an active opioid prescription is slightly smaller and no longer statistically significant. It is also possible that commanders do more drug testing in response to soldiers receiving more prescriptions/opioid prescriptions. If there is a peer spillover from health or health behaviors this could also impact our estimates. Table A8 column (5) includes controls for the percent of soldiers in the company that received a prescription, received an opioid prescription, or had a health rating in the four months prior to the start of the commander's command. Controlling for peer health does not impact our results. Column (6) shows the results of including controls for peer drug and alcohol use and peer health as well as commander demographics and quality discussed in the previous section. The estimates for illicit drug use and opioid prescriptions are almost identical to our main specification. This suggests that our results are likely due to an increase in exposure to drug testing not correlated commander characteristics or peer spillovers.

An alternative method to address concerns over the relationship between commander testing frequency and peer drug use is to estimate commander testing frequency controlling for peer characteristics.[25] We do this by regressing a dummy variable for the number of inspection tests a company has in a month on a fixed effect for each commander controlling for the racial, gender, educational, and rank composition of the company, the average age and AFQT score of soldiers in the company, and the percent of soldiers that are married or have kids. Importantly we also control for the percent of soldiers in the company who needed a waiver to enlist in the military, an indicator for prior criminal history or drug use, as well as the number of soldiers in the current month, previous month,

25. This is similar calculating a risk adjusted hospital quality.

and previous quarter with a positive drug test or drug or alcohol criminal offense.

$$NumTestDays_{c,t} = CDR_j^{radj} + V_{c,t} + \epsilon_{c,t}$$

We will define the fixed effect, CDR_j^{radj}, as the risk adjusted commander testing frequency.

Before looking the impact of the risk adjusted commander testing frequency on soldier outcomes, we examine the correlation between the risk adjusted and original measure of commander testing frequency. Regressing CDR_j^{radj} on CDR_j gives a coefficient of 0.986 (std error 0.00002) and and r-squared of 0.965. This is reassuring. The fact that the two measures of commander testing frequency are so closely related suggests that the frequency of inspection tests is not being driven by the characteristics of soldiers in the company. This is inline with the claim that inspection testing is driven by Army policy, company activities, and commander preferences, not soldier drug use. Figure 4 compares the impact of a one day increase in risk adjusted commander testing frequency to our main estimates. There is no discernible difference. This further supports the assumption that our findings of a decrease in illicit drug use and an increase in prescription opioid use and alcohol abuse are driven by the increase in testing, not correlated peer spillovers.

7.3 Heterogeneity Based on the Cost of Using Drugs

One testable implication of our model of drug use proposed in section 6 (eq 1) is that individuals with a higher cost to being discharged from the Army for drug use should be more responsive to increases in the likelihood of being drug tested. While we cannot identify individual's cost of being discharged, we can identify groups that likely have greater costs of separating from the Army. The two dimensions of heterogeneity that we explore are race and family status.

Previous work has shown that volunteering to enlist in the Army increases both short-term and long-run wages of black individuals while having a smaller to null effect on the wages of white individuals (Greenberg et al. 2022). The authors posit that the increased wages for black individuals are due to increase opportunity while in the Army and Army service serving as a positive signal to civilian employers that may overcome racial bias. While this does not speak directly to our question of the cost of being discharged for drug use, it seems reasonable to assume that the signaling value of Army service is much smaller or potentially negative if an individual is discharged for misconduct. Therefore, if Army service is a more valuable signal for black soldiers than white soldiers, the cost of being discharged for drug use would be greater for black soldiers. To test this we add an interaction term to equation 5.2 interacting $CDR_{j(i)}$ and a dummy variable for soldier i being black. We restrict the sample to soldiers who are white or black as Greenberg et al. (2022) does not show conclusive results for Hispanic or other race individuals. Table A9 panel A shows the results. Black soldiers experience a much greater decline in drug use when assigned to a high testing commander than white soldiers. Explicitly, a one day per month increase in average commander testing frequency decreases the likelihood that a black solider uses drugs by 0.116pp (mean 0.562%) compared to 0.034 (mean 0.317%) for whites soldiers. Most of this difference comes from black soldiers being much less likely to use marijuana when assigned to a higher testing commander relative to white soldiers assigned to a higher testing commander. Black soldiers are also more likely to see a decrease in opiate use, but the difference is not statistically significant. This is inline with the prediction from our model.

There is also reason to believe that soldiers who are married or have children have a higher cost of being discharged for misconduct. There are two reasons for this. First, soldiers with dependents receive an additional approximately $200-$1000 per month as part of their basic allowance for housing (BAH). Additionally, some Army benefits, including healthcare and educational benefits, are more valuable for individuals with family members who can use the benefits. For these reasons, we argue that our model predicts that soldiers with dependents will see a larger decrease in drug use when assigned to a higher testing commander. Table A9 Panel B shows the results of interacting commander testing frequency with a dummy variable for whether the soldier currently has a spouse or dependent child. While we lack statistical significance, soldiers with dependents see a much greater decrease in drug use when assigned to high testing commanders than soldiers without dependents. Being assigned to a commander that averages 1 more test day per month decrease drug use among soldiers with dependents by 0.065pp compared 0.025pp for soldiers without dependents. Interesting, this difference is driven by soldiers with dependents having greater decreases in marijuana and cocaine use and a potentially smaller decrease in opiate use. While we lack statistical precision, the fact that black soldiers and soldiers with dependents appear to be more responsive to increases in the drug testing suggests that soldiers who gain greater benefits from Army service and for whom being discharged for drug use is more costly

8 Robustness: Estimating Commander Testing Frequency

Throughout the paper we have measured commander testing frequency as the average number of days with an inspection test in the four months prior to the soldier arriving at the company. There is no a priori reason for choosing to take this average over four months. In fact, using such a small number of months may introduce a significant amount of noise into our measure. We can do an empirical bayes exercise to attempt to reduce the noise. Alternatively we can calculate the commander testing frequency using more months. This approach has the drawback of reducing our effective sample size. For a soldier to be in our sample, they need to arrive in the fifth month of their commander's command or later so that we can estimate the commander testing over the previous four months of the command. If we were to use, for example, the previous eight months of the command to calculate the testing frequency, only soldiers who arrived in the ninth month of the command or later would be included in the sample.

8.1 Empirical Bayes Shrinkage

One potential concern is that variation in the estimated commander testing frequency is being driven by noise and not an underlying characteristic of the commander. While the relatively clear linear relationship between commander testing frequency and the likelihood that soldiers are tested in the future shown in figure 2 panel (a) indicates that the variation is not noise, we can construct an Empirical Bayes (EB) version of our fixed effects. Following the linear shrinkage approach of Walters (2024), we construct the following estimate.

$$CDR^{EB}_{j(i)} = (\frac{\hat{\sigma}^2_{CDR}}{\hat{\sigma}^2_{CDR} + s^2_{j(i)}})\hat{CDR}_{j(i)} + (1 - \frac{\hat{\sigma}^2_{CDR}}{\hat{\sigma}^2_{CDR} + s^2_{j(i)}})\hat{\mu}_{CDR}$$

Here, μ_{CDR} and σ^2_{CDR} are the mean and variance of the estimated commander testing frequency while $s_{j(i)}$ is the standard deviation of each commander. If the variance of the fixed effects is large relative to the standard error of each fixed effect, the EB estimate will be close to the original fixed effect. If the standard error is large relative to the variance, the EB estimate will be "shrunk" towards the mean of the estimated fixed effects. Figure A4 shows how the EB testing frequency compares to the original testing frequency measure. Figure 4 panel (a) shows how the EB estimates compare to the original estimates. The EB estimates are very similar to the estimates using our primary measure.

8.2 Alternative Testing Frequency Periods

As stated in the beginning of this section, increasing the length of time used to calculate the commander testing frequency could increase the reliability of our commander testing frequency measure to predict a soldier's future likelihood of being tested. Using a longer period has the drawback of reducing the sample size. Additionally, if a commander's recent testing behavior is more relevant than farther back testing behavior, extending the period used to calculate commander testing frequency could reduce our ability to predict a soldier's future likelihood of being tested. Figure 4 panel (b) shows the results when calculating commander testing frequency using the four, six, eight, or ten previous months. As we use more months to calculate the commander testing frequency the estimated impact of commander testing frequency on the likelihood of being tested in the future increases. This suggests that using more months makes commander testing frequency a less noisy measure of the probability of a soldier being tested. The impact on illicit drug use and prescription opioid use are fairly consistent across different specifications. Note that increasing the months used to calculate commander testing frequency decreases our sample size.

9 Conclusion

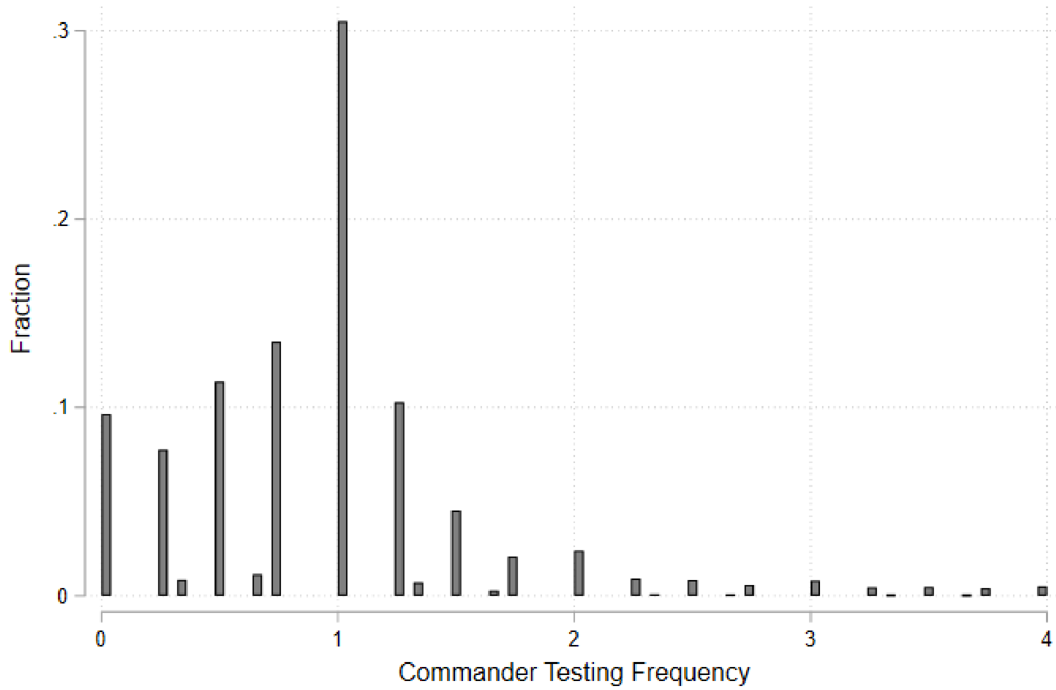
We have shown that soldiers assigned to commanders who conduct more frequent drug tests are less likely to use illicit drugs and more likely to receive an opioid prescription. Being assigned to a commander who averages 1 additional testing day per month increases the likelihood that a soldier is drug tested in the future by 12%. This decreases the likelihood that the soldier tests positive for an illicit drug conditional on being tested by 12% while increasing the likelihood of having an active opioid prescription by 3%. The increase in opioid prescriptions offsets about 78% of the decrease in illicit opioid use.

We propose a simple deterrence model to analyze the impact of drug testing. Our model predicts that drug testing will have a larger impact on decreasing the use of a specific substance if that substance is detectable for longer or has close legal substitutes. We find the decrease in illicit drug use to be evenly split between marijuana, which remains detectable in urine samples for longer than other tested for substances, and illicit opioids, which have direct prescription substitutes. We find drug testing to have little to no impact on the use of cocaine, a substance that is metabolized quickly and as no close legal substitutes in our setting. Our model also predicts that individuals for whom being caught with drugs and kicked out of the Army is more costly should be more responsive to increased drug testing. We find suggestive evidence that black soldiers and soldiers with dependents, two groups that may receive greater relative benefits to Army service, have greater decreases in drug use when assigned to a high testing commander compared to white soldiers and soldiers without dependents respectively.

We provide the caveat that commander testing frequency is correlated with measures of both commander quality and peer alcohol and drug use. Our estimates, however, are robust to controlling for these measures suggesting that the results are driven by an increased probability in being tested not other commander or company characteristics. We highlight the importance of accounting for possible substitution effects when analyzing the effect of drug testing policies. Depending on the setting, much of the benefit in a decrease in illicit drug use may be offset if individuals replace illicit drug use with prescription opioid use.

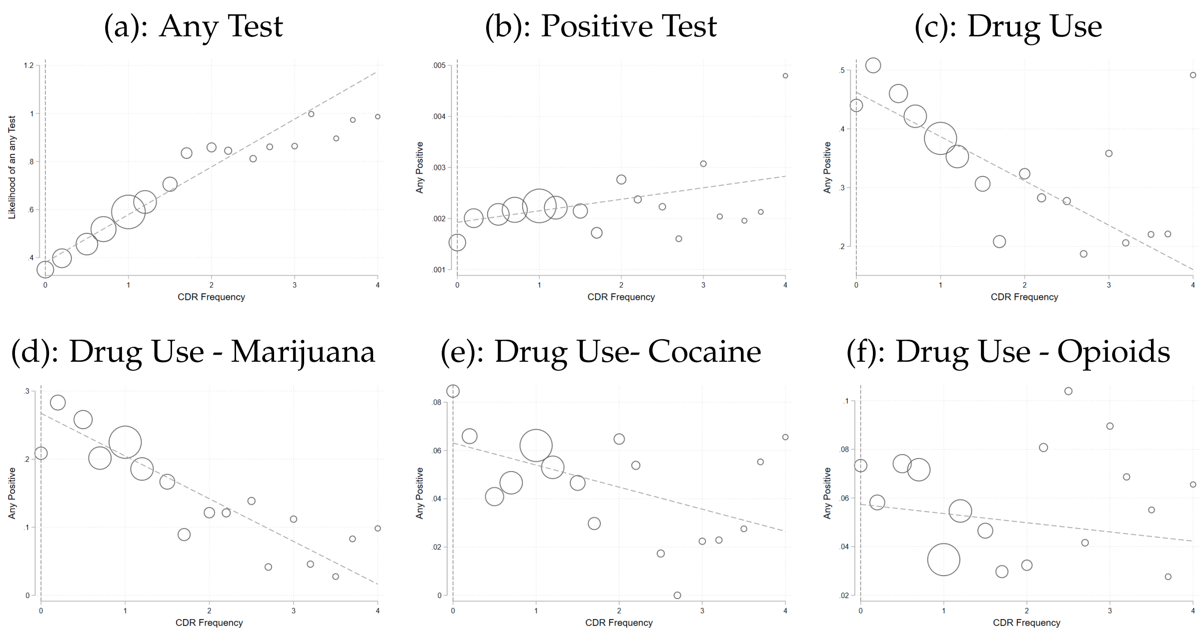
Figures

Figure 1: Distribution of Commander Testing Measures



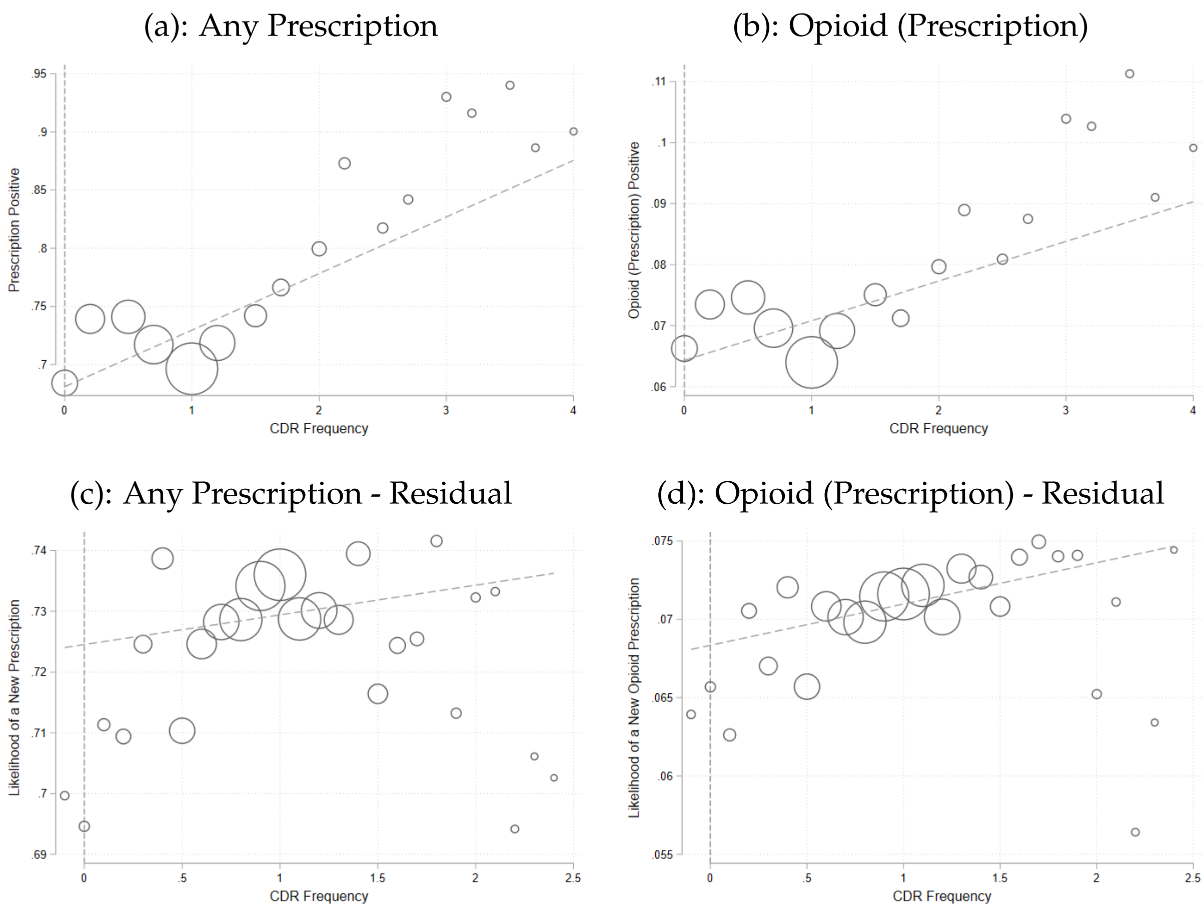
Notes: The figure show the distribution of the commander testing frequency fixed effect. We calculate the average number of testing days per month over the previous four months for 7,470 company commanders, for each month of the command starting in month 5. The plurality of commanders conduct 1 inspection test per month.

Figure 2: Relationship between Commander Testing and Likelihood of being Tested and Testing Positive



Notes: This figure shows the relationship between the estimated commander testing frequency fixed effect and the likelihood that a soldier is drug tested (panel a) tests positive (panel b), and tests positive conditional on being tested (panel c-f). Note that Panel b reflects both the increase in the likelihood of testing positive due to the increased likelihood in being tested as well as any change in the likelihood of using drugs. Under the assumption that inspection drug tests are independent of soldier drug use, Panels c-f show represent the impact of commander testing frequency on soldier drug use. Figure A2 shows the relationship between commander testing frequency and drug use after controlling for the fixed effects in our main specification.

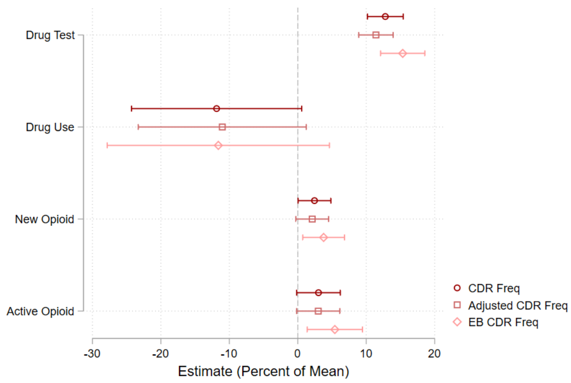
Figure 3: Relationship between Testing Frequency and Prescriptions



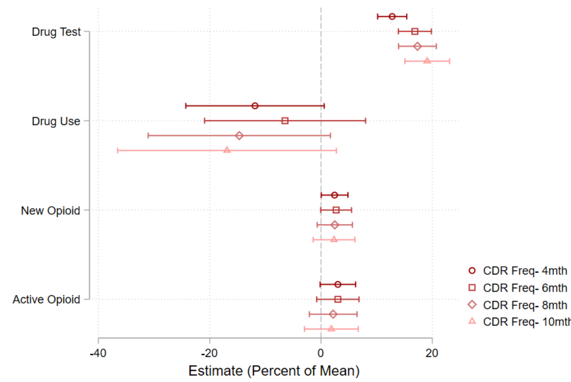
Notes: This figure shows the relationship between the estimated commander testing frequency fixed effect and the likelihood of a receiving a new prescription. Panels (c) and (d) show the residual relationship between the estimated commander testing frequency fixed effect and the likelihood of a soldier receiving a new prescription after controlling for the interaction of soldier brigade, occupation, rank, and quarter of arrival at the company plus a separate fixed effect for year-month. This illustrates the variation being used in our main specification. To create this plot we regress commander testing frequency and each outcome on the fixed effects. We then take the residual from each regression and add back the sample mean.

Figure 4: Robustness to Alternative Measures of Commander Testing Frequency

(a): Alternative CDR Frequency



(b): Alternative Time Periods



Notes: This figure shows the estimates and 95% confidence intervals from regressing each outcome on commander testing frequency. Each estimate is scaled by the outcome mean. Every regression controls for soldier characteristics and includes fixed effects for the interaction of brigade, occupation, rank, and arrival quarter and a separate fixed effect for calendar year-month. This is the same specification as column (3) of table 3. Panel (a) compares estimates using our primary measure of commander testing frequency to estimates using the risk adjusted and empirical bayes measures. Panel (b) shows how the estimates change as we calulcate the commander testing frequency using the previous 6, 8, and 10 months of the command compared to our primary measure which uses the first 4 months. Note that when we use the 10 months to calculate the commander testing frequency we drop 28% of our sample

Tables

Table 1: Balance Table: Soldier Characteristics are Uncorrelated with Commander Testing Frequency and the Likelihood of having an Inspection Test

Panel A: Commander Testing Frequency

	Female	Black	Hispanic	Other Race	Married	Has Kids	HSG+	Age	AFQT	Moral Waiver
Commander Testing Frequency	-0.0672	0.4166***	-0.3159**	-0.0023	0.2819	0.1106	-0.0263	0.0359	0.0858	-0.0004
	(0.1238)	(0.1532)	(0.1342)	(0.0925)	(0.1748)	(0.1493)	(0.0814)	(0.0261)	(0.0611)	(0.0005)
Covariate Mean	14.788	24.643	18.004	7.052	39.404	27.598	94.127	21.205	55.512	0.024
Observations	375234	375234	375234	375234	375234	375234	375234	375230	374382	375234
P-value on Joint Test	0.010									

Panel B: Inspection Tests

	Female	Black	Hispanic	Other Race	Married	Has Kids	HSG+	Age	AFQT	Moral Waiver
any_t_test	-0.1669***	-0.1158***	0.0120	0.0500*	-0.4683***	-0.3100***	-0.0061	-0.0067*	0.0421**	-0.0205
	(0.0360)	(0.0437)	(0.0435)	(0.0273)	(0.0509)	(0.0427)	(0.0262)	(0.0039)	(0.0176)	(0.0164)
Covariate Mean	15.652	26.081	17.858	7.213	45.467	33.703	94.488	21.337	55.681	2.523
Observations	75862666	75862666	75862666	75862666	75862666	75862666	75862666	75861179	75680017	75862666
P-value on Joint Test	0.000									

Panel C: Inspection Tests - Female and Dependent Controls

	Black	Hispanic	Other Race	HSG+	Age	AFQT	Moral Waiver
Inspection Drug Tests	-0.1073**	0.0246	0.0461*	-0.0127	0.0021	0.0292*	-0.0173
	(0.0437)	(0.0435)	(0.0273)	(0.0262)	(0.0038)	(0.0176)	(0.0164)
Covariate Mean	26.081	17.858	7.213	94.488	21.337	55.681	2.523
Observations	75862666	75862666	75862666	75862666	75861179	75680017	75862666
P-value on Joint Test	0.120						

Notes: Panel A shows the results of regressing each soldier characteristic on the estimated commander testing frequency and the number of inspection tests the soldier has including fixed effects for the interaction of the soldier's occupation, grade, location, and quarter of arrival as well as the calendar year-month (equation 4). Panel A keeps the first observation for each soldier at each company. This is a test for random assignment of soldiers to commander testing. Panels B and C shows the results of regressing soldier characteristics on a dummy variable for having an Inspection drug tests including the same set of fixed effects as panel A. Panel C adds controls for soldier gender and whether the soldier has any dependents. This supports the claim that Inspection tests, including Inspection Random, Inspection Unit, and Inspection Other tests, do not target certain individuals as well as the claim that the likelihood of having an Inspection test is independent of the soldiers drug use. $* \ p < 0.1$, $^{**} \ p < 0.05$, $^{***} \ p < 0.01$

Table 2: Correlation between Company Characteristics and Commander Testing Frequency

Panel A: Commander Characteristics

	Female	Black	Hispanic	Other Race	Married	Age	Graduate Degree	Years of Service
Commander Testing Frequency	0.0520	0.0301	-1.1209**	-0.2254	0.9281	-0.0100	0.3387	-0.0832
	(0.6338)	(0.5538)	(0.5170)	(0.5179)	(0.7875)	(0.0715)	(0.6747)	(0.0599)
Covariate Mean	13.208	10.589	8.643	8.259	72.106	31.384	19.624	8.167
Observations	85922	85922	85922	85922	85922	85922	85922	85922
P-value on Joint Test	0.160							

Panel B: Commander Quality Measures

	SAT	USMA	Petersen Competitive	OER MQ %	OER MQ % (Future)	BZ Major
Commander Testing Frequency	-4.2120	0.2590	0.3804	2.1629***	1.4314***	1.2068**
	(3.8452)	(0.6670)	(0.4892)	(0.6866)	(0.5043)	(0.5431)
Covariate Mean	1166.878	20.043	9.510	51.754	55.255	5.439
Observations	49889	85922	85922	85481	85489	55928
P-value on Joint Test	0.000					

Panel C: Company Characteristics determined prior to the start of the command

	Inspection Test Frequency	Positive Drug Test	Alcohol Criminal Offense	Any Prescription	Opioid-Prescription	Health Rating
Commander Testing Frequency	0.1767***	0.0210***	0.0186**	0.3078**	0.0276	0.0001
	(0.0145)	(0.0050)	(0.0075)	(0.1345)	(0.0228)	(0.0005)
Covariate Mean	0.877	0.148	0.097	17.100	2.120	0.498
Observations	85922	85922	85922	85922	85922	85922
P-value on Joint Test	0.000					

Notes: This table shows the results of regressing each company/commander characteristic on commander testing frequency. We keep one observation per commander per month. Each regression includes fixed effects for the interaction of Brigade, commander occupation, and year-quarter to mimic our main specification. High testing commander are less likely to be Hispanic. Commander testing frequency is positively correlated with the number of soldiers who have a prescription or opioid prescription. High testing commanders are more likely to be in charge of companies who experienced higher drug testing frequency, number of positive drug tests, and alcohol criminal offenses prior to the commander's arrival. A8 describes how this impacts our results. $* \, p < 0.1$, $^{**} \, p < 0.05$, $^{***} \, p < 0.01$

Table 3: Impact of Testing Frequency of Likelihood of being Tested and Using Drugs

	(1)	(2)	(3)
	Panel A: Likelihood of Being Tested (Inspection Test)		
Commander Testing Frequency	0.1384***	0.0732***	0.0729***
	(0.0072)	(0.0076)	(0.0076)
Observations	74,393,215	74,393,211	74,212,774
Dep Var Mean	0.570	0.570	0.570
R^2	0.003	0.004	0.004
	Panel B: Likelihood of Using Drugs		
Commander Testing Frequency	-0.0364*	-0.0414*	-0.0428*
	(0.0191)	(0.0229)	(0.0229)
Observations	406,462	393,527	392,705
Dep Var Mean	0.372	0.360	0.361
R^2	0.094	0.139	0.139
FE: Calendar Year-Month	X	X	X
FE: Post x Occupation x Rank x Arrival Qtr	X		
FE: Brigade x Occupation x Rank x Arrival Qtr		X	X
Individual Characteristics			X

Notes: This table shows the impact of a one-unit increase in commander testing frequency (one additional testing day per month) on the likelihood that the soldier is tested and uses drugs. Panel A looks at the impact on the likelihood of a soldier having an inspection test on a given day. Panel B looks at the likelihood of a soldier testing positive in an inspection test conditional on having at least one inspection test. Column (1) includes fixed effects for the interaction of soldier location, occupation, rank, and quarter of arrival at the company. This is represents the factors used by the Army to assign soldiers to units. Column (2) uses brigade instead of location. This is meant to control for variation in daily assignments between soldiers in different brigades as this variation may also be correlated with the likelihood that companies conduct drug tests or soldiers seek medical attention. Column (3) controls for individual characteristics including race, gender, marriage status, education, having dependents, AFQT score, and enlisting with a moral waiver. All standard errors are clustered at the commander level. $* p < 0.1$, $^{**} p < 0.05$, $^{***} p < 0.01$

Table 4: Impact of Testing Frequency on Drug Use (Positive Inspection Tests)

	(1) Any Drug	(2) Marijuana	(3) Cocaine	(4) Opiates
Commander Testing Frequency	-0.0428*	-0.0219	-0.0042	-0.0223**
	(0.0229)	(0.0134)	(0.0083)	(0.0106)
Observations	392,705	392,705	392,705	392,705
Dep Var Mean	0.361	0.189	0.051	0.051
R^2	0.139	0.138	0.107	0.176

Notes: This table shows the impact of a one-unit increase in commander testing frequency (one additional testing day per month) on the likelihood that the soldier uses drugs different types of drugs. Each column includes controls for soldier characteristics as well as fixed effects for the interaction of location, occupation, rank, and arrival quarter and a separate fixed effect for calendar year-month. This is the same specification as table 3 column (3). All standard errors are clustered at the commander level. $* \, p < 0.1$, $^{**} \, p < 0.05$, $^{***} \, p < 0.01$

Table 5: Impact of Commander Testing Frequency on Prescriptions

	(1)	(2)	(3)
	Panel A: Likelihood of Receiving a Prescription		
Commander Testing Frequency	0.0185***	0.0034	0.0044
	(0.0045)	(0.0049)	(0.0048)
Observations	58,155,018	58.155,015	58,005,488
Dep Var Mean	0.932	0.932	0.931
R^2	0.005	0.008	0.008
	Panel B: Likelihood of Receiving an Opioid Prescription		
Commander Testing Frequency	0.0035***	0.0022*	0.0022**
	(0.0010)	(0.0011)	(0.0011)
Observations	58,155,018	58,155,015	58,005,488
Dep Var Mean	0.090	0.090	0.090
R^2	0.003	0.004	0.004
FE: Calendar Year-Month	X	X	X
FE: Post x Occupation x Rank x Arrival Qtr	X		
FE: Brigade x Occupation x Rank x Arrival Qtr		X	X
Individual Characteristics			X

Notes: This table shows the impact of a one unit increase in commander testing frequency on the likelihood of a soldier receiving a prescription. Panel A looks at the impact on having any new prescription. Panel B looks at the likelihood of a soldier receiving an opioid prescription. As in table 3 column (1) includes fixed effects for the interaction of soldier location, occupation, rank, and quarter of arrival at the company. This is represents the factors used by the Army to assign soldiers to units. Column (2) uses brigade instead of location. This is meant to control for variation in daily assignments between soldiers in different brigades as this variation may also be correlated with the likelihood that companies conduct drug tests or soldiers seek medical attention. Column (3) controls for individual characteristics including race, gender, marriage status, education, having dependents, AFQT score, and enlisting with a moral waiver. All standard errors are clustered at the commander level. $* \ p < 0.1$, $^{**} \ p < 0.05$, $^{***} \ p < 0.01$

References

Alpert, Abby, David Powell, and Rosalie Liccardo Pacula. 2018. "Supply-Side Drug Policy in the Presence of Substitutes: Evidence from the Introduction of Abuse-Deterrent Opioids." *American Economic Journal: Economic Policy* 10 (4): 1–35. https://doi.org/10.1257/pol.20170082.

Army Substance Abuse Program. 2018. *Unit Prevention Leader's Urinalysis Collection Handbook.* Washington D.C.

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

Bruhn, Jesse, Kyle Greenberg, Matthew Gudgeon, Evan K Rose, and Yotam Shem-Tov. 2024. "The effects of combat deployments on veterans' outcomes." *Journal of Political Economy* 132 (8): 000–000.

Carpenter, Christopher S. 2006. "Workplace Drug Testing and Worker Drug Use." *Health Services Research* 42 (2): 795–810.

Carrieri, Vincenzo, Leonardo Madio, and Francesco Principe. 2020. "Do-It-Yourself medicine? The impact of light cannabis liberalization on prescription drugs." *Journal of Health Economics* 74.

Department of the Army. 2016. *Army Regulation 600-85: The Army Substance Abuse Program.* Washington D.C.

———. 2019. *Army Regulation 614-200: Enlisted Assignments and Utilization Management.* Washington D.C.

Ghimire, Ramesh, Curtis Florence, and Gery Guy. 2025. "Productivity Losses From Substance Use Disorder in the U.S. in 2023." *American Journal of Preventive Medicine*, https://doi.org/https://doi.org/10.1016/j.amepre.2025.108102.

Greenberg, Kyle, Matthew Gudgeon, Adam Isen, Corbin Miller, and Richard Patterson. 2022. "Army service in the all-volunteer era." *The Quarterly Journal of Economics* 137 (4): 2363–2418.

House, The White. 2025. *The Staggering Cost of the Illicit opioid Epidemic in the United States.* https://www.whitehouse.gov/articles/2025/03/the-staggering-cost-of-the-illicit-opioid-epidemic-in-the-united-states/. Last accessed: December 15, 2025.

Lee, Doohee, and Michael W. Ross. 2011. "Management of Human Resources Associated with Misuse of Prescription Drugs: Analysis of a National Survey." *Journal of Health and human Servies Administration* 34 (2): 182–205.

Lucas, Philippe, and Zach Walsh. 2017. "Medical cannabis access, use, and substitution for prescription opioids and other substances: A survey of authorized medical cannabis patients." *International Journal of Drug Policy* 42:30–35.

Lucas, Philippe, Zach Walsh, Kim Crosby, Robert Callaway, Lynne Belle-Isle, Robert Kay, and Rielle Capler. 2016. "Substituting cannabis for prescription drugs, alcohol and other substances among medical cannabis patients: The impact of contextual factors." *Drug and Alcohol Review* 35 (3): 326–333.

Lyle, David S. 2006. "Using Military Deployments and Job Assignments to Estimate the Effect of Parental Absences and Household Relocations on Children's Academic Achievement." *Journal of Labor Economics* 24 (2): 319–350.

Moeller, Karen E., Kelly C. Lee, and Julie C. Kissack. 2008. "Urine Drug Screening: Practical Guide for Clinicians." *Mayo Clinic Proceedings* 83 (1): 66–67. https://doi.org/https://doi.org/10.4065/83.1.66.

Perrone, Dina, Randi D. Helgesen, and Ryan G. Fischer. 2013. "United States drug prohibition and legal highs: How drug testing may lead cannabis users to Spice." *Drugs: Education, Prevention and Policy* 20 (3): 216–224.

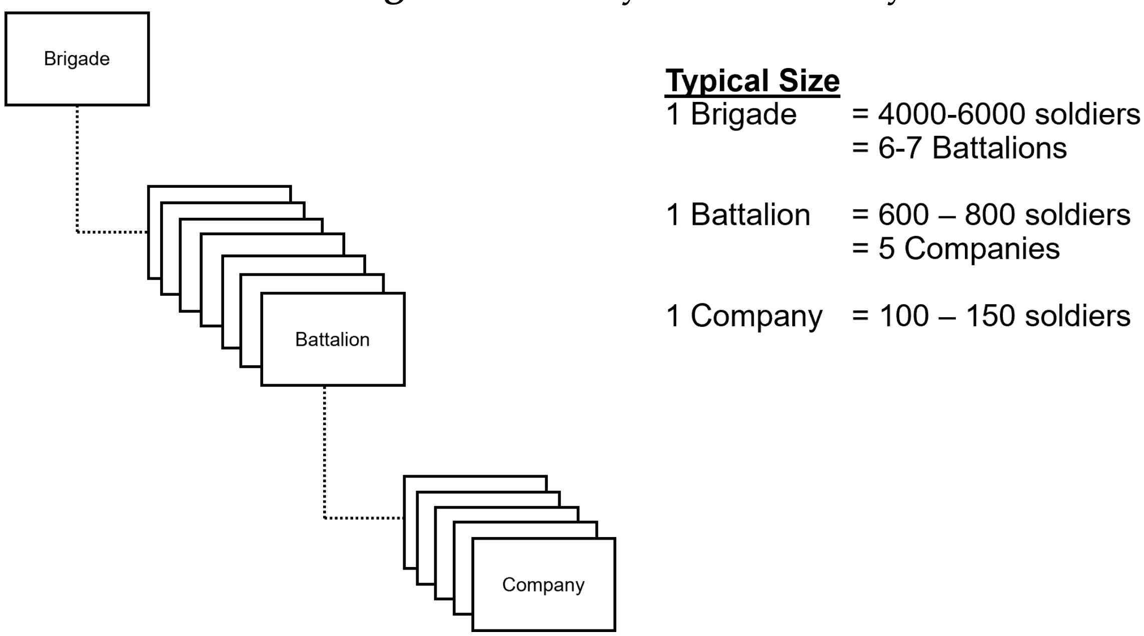
Services, EBSCO Information. 2024. *Economic Impact of Addiction.* https://www.ebsco.com/research-starters/economics/economic-impact-addiction. Last accessed: December 15, 2025.

Walters, Christopher. 2024. "Empirical Bayes methods in labor economics." Chap. 3 in *Handbook of Labor Economics,* 183–260. National Bureau of Economic Research. https://doi.org/https://doi.org/10.1016/bs.heslab.2024.11.001.

A Appendix

Appendix Figures

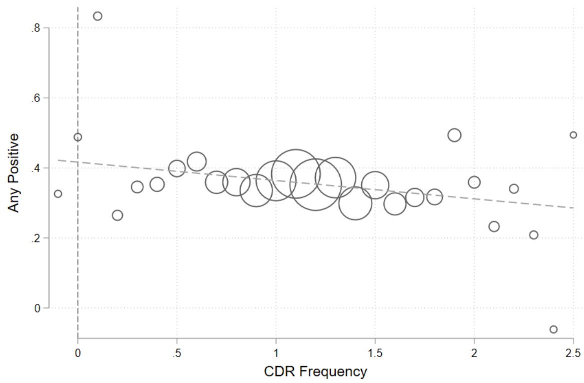
Figure A1: Army Unit Hierarchy



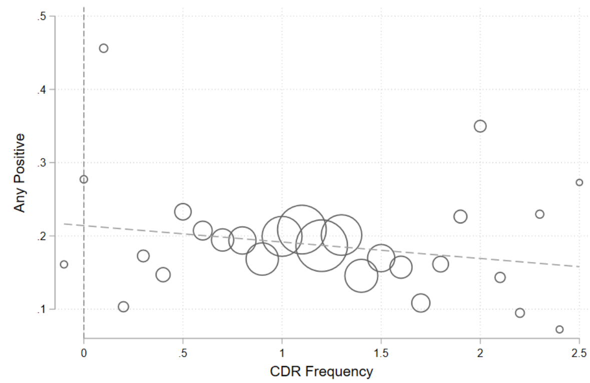
Notes: This figure displays the typical hierarchy and organization of units within an Army Brigade. Drug testing is managed and overseen within the Army at the battalion level, and implemented at the company level.

Figure A2: Relationship between Testing Frequency and Drug Use Rates

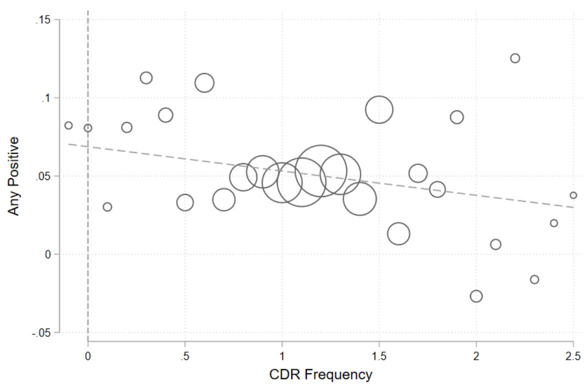
(a): Any Positive



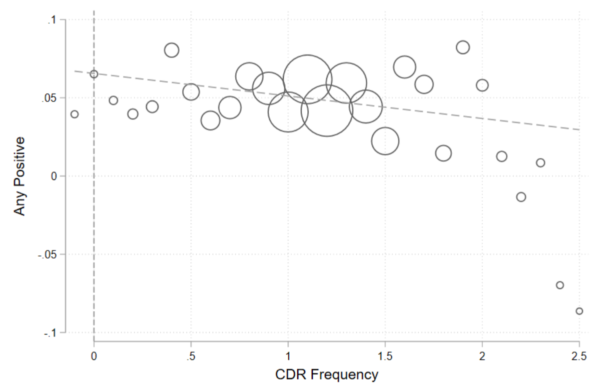
(b): Marijuana



(c): Cocaine

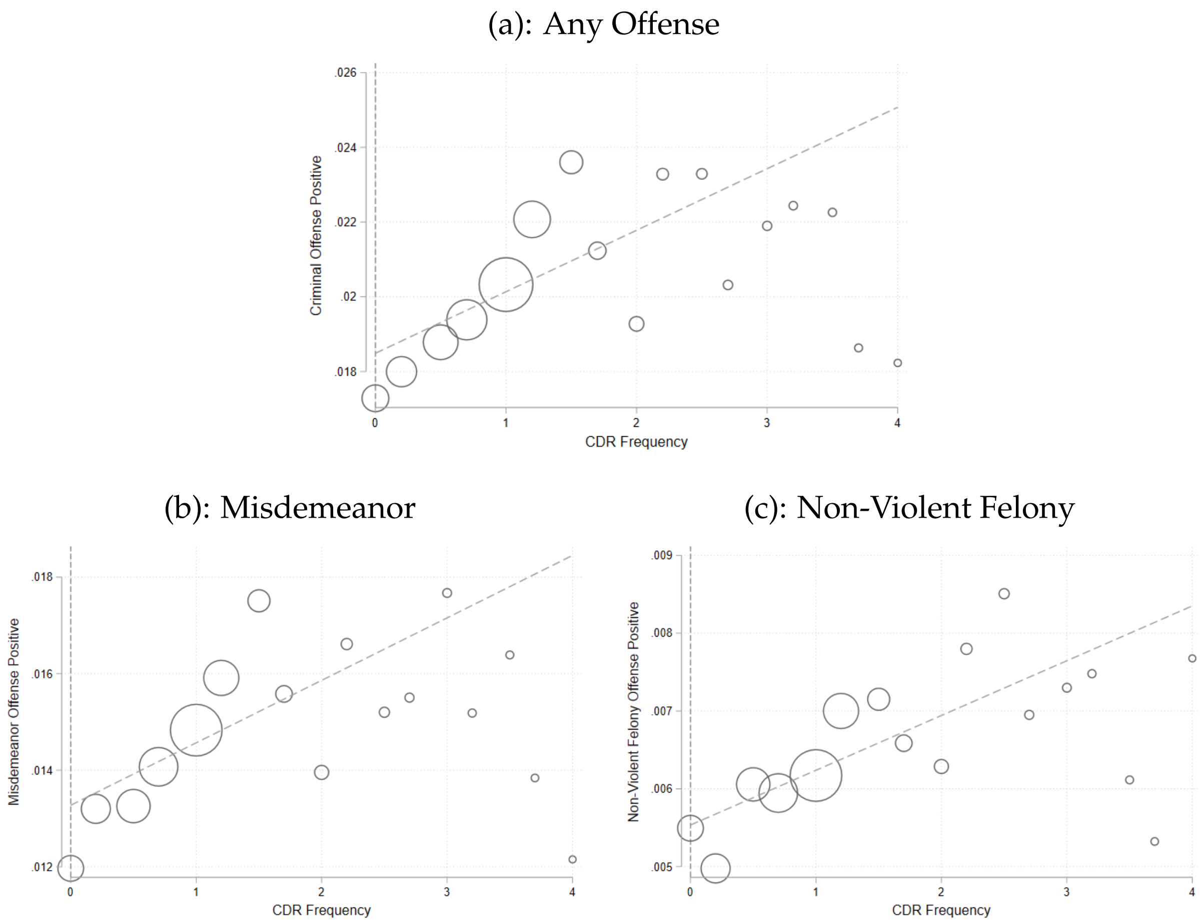


(d): Opioids



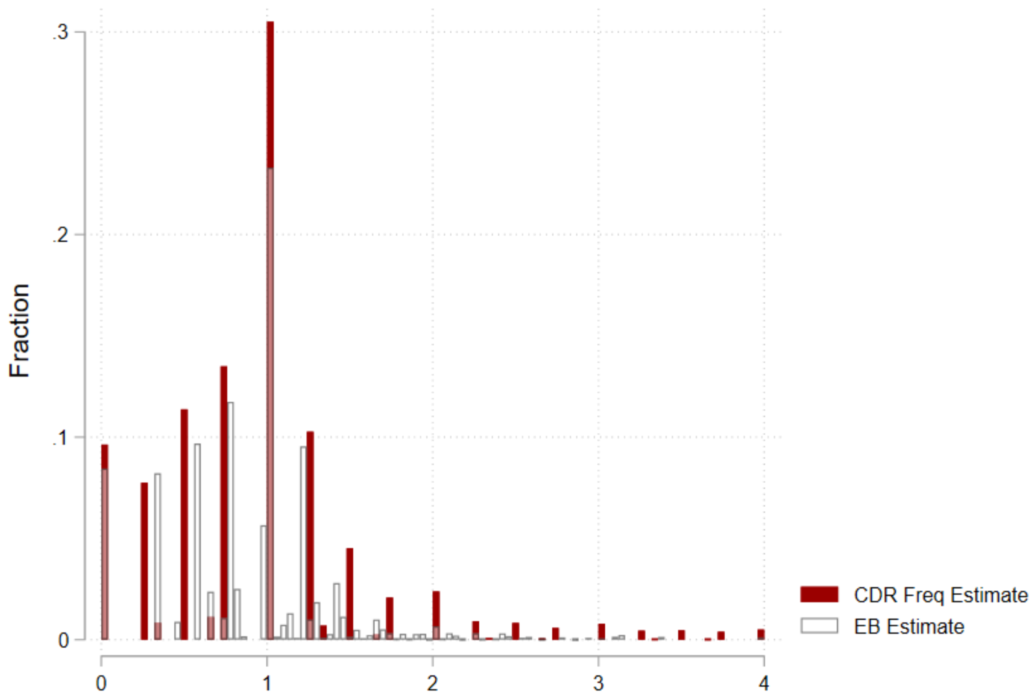
Notes: This figure shows the residual relationship between the estimated commander testing frequency fixed effect and the likelihood of a soldier testing positive in an Inspection test conditional on being tested after controlling for the interaction of soldier brigade, occupation, rank, and quarter of arrival at the company plus a separate fixed effect for year-month. This illustrates the variation being used in our main specification. To create this plot we regress commander testing frequency and each outcome on the fixed effects. We then take the residual from each regression and add back the sample mean.

Figure A3: Relationship between Testing Frequency and Criminal Offenses



Notes: This figure shows the relationship between the estimated commander testing frequency fixed effect and the likelihood of a soldier having a criminal offense.

Figure A4: Distribution of Commander Testing Measures



Notes: The red bars show the distribution of the commander testing frequency fixed effect. We calculate the average number of testing days per month over the previous four months for 7,470 company commanders, for each month of the command starting in month 5. The black outlines show the distribution of commander testing frequency after applying the linear shrinkage approach described in section 7.3.1.

Appendix Tables

Table A1: Drug Detection Duration

Drug	Time from Use to Positive Test	Positive Test Duration
Marijuana	1-3 hours	1-7 days
Crack/Cocaine	2-6 hours	2-3 days
Heroin/Opiates	2-6 hours	1-3 days
Amphetamine/Meth	4-6 hours	2-3 days
Angel Dust/PCP	4-6 hours	7-14 days
Ecstacy	2 to 7 hours	2 - 4 days
Benzodiazepine	2 -7 hours	1 - 4 days
Barbiturates	2 - 4 hours	1 - 3 weeks
Methadone	3 - 8 hours	1 - 3 days
Tricyclic Antidepressants	8 - 12 hours	2 - 7 days
Oxycodone	1 - 3 hours	1 - 2 days

Note: The above data comes from the Health Service Executive National Social Inclusion Office found at the following website: https://www.drugs.ie/drugs_info/about_drugs/how_long_do_drugs_stay_ in_your_system/

Table A2: Summary Statistics

Soldiers

	Mean
Female	0.15
White	0.50
Black	0.25
Hispanic	0.18
Other Race	0.07
Age	21.29
Married	0.39
Has Kids	0.28
AFQSC	56.39
Greater than High School Education	0.94
Years of Service	3.38
Drug Tests per Year	2.15
Fraction of Sample with Drug Charge	0.01
Observations	323,777

Commanders

	Mean
Female	0.13
White	0.73
Black	0.10
Hispanic	0.08
Other Race	0.08
Age	30.50
Married	0.69
Has Kids	0.44
Years of Service	7.29
Months in Command	15.98
Observations	7,488

Notes: This table shows summary statistics for each soldier and commander in out sample. Individual race and AFQSC is based on data from when the individual entered the Army. Data on age, family status, education, years of service, and prior drug charge come from the first month the individual is observed in the sample. Drug tests per year is an average across all soldiers in our sample's full tenure in the Army. Months in Command is the average length of command for each commander's assignment as a company commander in our sample.

Table A3: Balance Table: Soldier Characteristics are Uncorrelated with Commander Testing Frequency and the Likelihood of having an Inspection Test

	Black	Hispanic	Other Race	HSG+	Age	AFQT	Moral Waiver
Inspection Drug Test (IR/IU)	-0.1023**	0.0327	0.0394	-0.0076	0.0023	0.0248	-0.0163
	(0.0454)	(0.0454)	(0.0285)	(0.0271)	(0.0040)	(0.0184)	(0.0170)
Covariate Mean	26.081	17.858	7.213	94.488	21.337	55.681	2.523
Observations	75862666	75862666	75862666	75862666	75861179	75680017	75862666
P-value on Joint Test	0.260						

Notes: This table shows the results of regressing soldier characteristics on a dummy variable for having an Inspection Random or Inspection Unit drug test. Unlike in table 1 panel C, this table excludes Inspection Other (IO) tests. There may be concerns over the predictability of IO tests and the potential ability of commanders to selectively enforce policies requiring IO tests. This supports the claim that Inspection Random and Inspection Unit tests do not target certain individuals. This supports the claim that the likelihood of having an Inspection test is independent of the soldiers drug use. All standard errors are clustered at the commander level. $* p < 0.1$, $^{**} p < 0.05$, $^{***} p < 0.01$

Table A4: Impact of Testing Frequency on Drug Use (Positive IR/IU Tests)

	(1) Any Drug	(2) Marijuana	(3) Cocaine	(4) Opiates
Testing Frequency	-0.0490**	-0.0241*	-0.0060	-0.0255**
	(0.0240)	(0.0137)	(0.0087)	(0.0110)
Observations	369,477	369,477	369,477	369,477
Dep Var Mean	0.362	0.189	0.052	0.051
R^2	0.140	0.138	0.107	0.175

Notes: This table shows the impact of a one-unit increase in commander testing frequency (one additional testing day per month) on the likelihood that the soldier uses drugs different types of drugs. Each column includes controls for soldier characteristics as well as fixed effects for the interaction of location, occupation, rank, and arrival quarter and a separate fixed effect for calendar year-month. Unlike in table 4, we only use Inspection Random and Inspection Unit tests (excluding Inspection Other tests). This addresses potential concerns that soldiers who miss company IR or IU tests can predict that they will have an IO test soon and adjust their behavior. Soldiers test positive during IO tests at very similar rates to IR and IU tests, and restricting to IR and IU tests leads to slightly learger estimates. All standard errors are clustered at the commander level. $*\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table A5: Impact of a 1 std deviation increase in testing frequency on Opioid Prescriptions

	(1) Any Opioid	(2) Oxycodone	(3) Hydrocodone	(4) Codeine	(5) Fentanyl	(6) Morphine
	Panel A: Likelihood of Having a New Opioid Prescription					
Commander Testing Frequency	0.00224**	0.00208***	-0.00047	-0.00000	0.00001	0.00012**
	(0.00112)	(0.00077)	(0.00042)	(0.00025)	(0.00003)	(0.00006)
Observations	58,005,488	58,005,488	58,005,488	58,005,488	58,005,488	58,005,488
Dep Var Mean	0.0903	0.0401	0.0154	0.0057	0.0001	0.0003
R^2	0.0037	0.0032	0.0027	0.0025	0.0014	0.0027
	Panel B: Likelihood of having an Active Prescription Opioid					
Testing Frequency	0.01746*	0.00640	-0.00255	-0.00133	0.00045	0.00119**
	(0.00935)	(0.00453)	(0.00237)	(0.00166)	(0.00044)	(0.00050)
Observations	58005488	58005488	58005488	58,005,488	58005488	58,005,488
Dep Var Mean	0.5777	0.1952	0.0648	0.0282	0.0024	0.0020
R^2	0.0555	0.0319	0.0261	0.0271	0.0378	0.0438

Notes: This table shows the impact of a one-unit increase in commander testing frequency (one additional testing day per month) on the likelihood that the soldier has a opioid prescription. Panel A shows the impact on having a new prescription while pane B shows the impact on having an active prescription. For example, if a soldier is written a 7 day prescription on January 1st, the soldier has a new prescription on January 1st and an active prescription for January 1st-7th. Each regression controls for individual characteristics include race, gender, marriage status, education, having dependents, AFQT score, and enlisting with a moral waiver and fixed effects for the interaction of brigade, occupation, rank, and quarter of arrival as well as a separate fixed effect for calendar year-month.

Table A6: Impact of Commander Testing Frequency on Overall Opioid Use

	(1) Illicit Opioid	(2) Active Prescription Opioid
Commander Testing Frequency	-0.0223**	0.0175*
	(0.0106)	(0.0094)
Observations	392,705	58,005,488
Dep Var Mean	0.051	0.578
R^2	0.176	0.055

Notes: This table compares the estimates of a 1 unit increase in commander testing frequency on illicit opioid use and prescription opioid use. Both columns repeat our main specification including controls for individual demographics and fixed effects for the year-month and the interaction of soldier brigade, occupation, rank, and quarter of arrival at the company. Column (1) shows the impact on the likelihood of testing positive for an opioid conditional on being tested (Table 3 Panel B, column 3). Column (2) shows the impact on the likelihood of having an active prescription. All standard errors are clustered at the commander level. $* \, p < 0.1$, $^{**} \, p < 0.05$, $^{***} \, p < 0.01$

Table A7: Impact of Testing Frequency on Criminal Offenses

	(1)	(2)	(3)
	Panel A: Likelihood of a Criminal Offense		
Commander Testing Frequency	0.0005	0.0005	0.0004
	(0.0004)	(0.0004)	(0.0004)
Observations	74,393,215	74,393,211	74,212,774
Dep Var Mean	0.020	0.020	0.020
R^2	0.001	0.002	0.002
	Panel C: Likelihood of a Misdemeanor Offense		
Commander Testing Frequency	0.0007**	0.0007**	0.0007**
	(0.0003)	(0.0003)	(0.0003)
Observations	74,393,215	74,393,211	74,212,774
Dep Var Mean	0.014	0.015	0.014
R^2	0.001	0.002	0.004
	Panel B: Likelihood of a Nonviolent Felony Offense		
Commander Testing Frequency	-0.0001	-0.0000	-0.0000
	(0.0002)	(0.0002)	(0.0002)
Observations	74,393,215	74,393,211	74,212,774
Dep Var Mean	0.006	0.006	0.006
R^2	0.001	0.003	0.003
FE: Calendar Year-Month	X	X	X
FE: Post x Occupation x Rank x Arrival Qtr	X		
FE: Brigade x Occupation x Rank x Arrival Qtr		X	X
Individual Characteristics			X

Notes: This table shows the impact of a one-unit increase in commander testing frequency (one additional testing day per month) on the likelihood that the soldier is tested and uses drugs. Panel A looks at the impact on the likelihood of a soldier having any criminal offense on a given day. Panels B and C looks at the likelihood of a soldier having a misdemeanor or nonviolent felony offense respectively. Notably 38% of all misdemeanors are alcohol related. Drug crimes make up about 40% of nonviolent felonies; however, to avoid a mechanical relationship between increased drug testing and increased drug crimes, we drop all drug crimes related to a urine test. Column (1) includes fixed effects for the interaction of soldier location, occupation, rank, and quarter of arrival at the company. This is represents the factors used by the Army to assign soldiers to units. Column (2) uses brigade instead of location. This is meant to control for variation in daily assignments between soldiers in different brigades as this variation may also be correlated with the likelihood that companies conduct drug tests or soldiers seek medical attention. Column (3) controls for individual characteristics including race, gender, marriage status, education, having dependents, AFQT score, and enlisting with a moral waiver. All standard errors are clustered at the commander level. $* p < 0.1$, $^{**} p < 0.05$, $^{***} p < 0.01$

Table A8: Importance of Controlling for Commander and Company Characteristics

	Main Specification (1)	Cdr Demographics (2)	Cdr Quality (3)	Company Drug/Alcohol (4)	Company Health (5)	All (6)
			Panel A: Inspection Drug Tests			
Commander Testing Frequency	0.0729***	0.0725***	0.0713***	0.0701***	0.0733***	0.0679***
	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0076)
			Panel B: Drug Use			
Commander Testing Frequency	-0.0428*	-0.0433*	-0.0420*	-0.0444*	-0.0425*	-0.0431*
	(0.0229)	(0.0230)	(0.0230)	(0.0230)	(0.0230)	(0.0232)
			Panel C: New Opioid Prescription			
Commander Testing Frequency	0.0022**	0.0022**	0.0023**	0.0022*	0.0022**	0.0023**
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
			Panel D: Active Opioid Prescription			
Commander Testing Frequency	0.0175*	0.0171*	0.0188**	0.0147	0.0170*	0.0162*
	(0.0094)	(0.0093)	(0.0094)	(0.0095)	(0.0094)	(0.0094)
Commander Demographics		X				X
Commander Quality			X			X
Company Drug Testing and Alcohol				X		X
Company RX and Health					X	X

Notes: This table shows how our key results change as we control for commander and company characteristics. Column (1) repeats our main specification including controls for individual demographics and fixed effects for the year-month and the interaction of soldier brigade, occupation, rank, and quarter of arrival at the company. Column (2) includes controls for commander characteristics including race, gender, education, age, and years of service. Column (3) includes measures of commander quality including their SAT score, the quality of their undergraduate institution, the percent of their past and future evaluations receiving ”Most Qualified”, and whether they were promoted to Major below the zone (ahead of schedule), Columns (4) and (5) include the average characteristics of company in the 4 months prior to the start of the commander. This can be primarily thought of as capturing for peer spillovers; however, these controls may also include impacts of the previous commander's actions. Column (4) includes the company monthly average number of inspection drug test days, monthly average of positive inspection drug tests, and monthly average number of alcohol criminal offenses. Column (5) includes the company monthly average number of soldiers with any prescription, an opioid prescription, or a service limiting health condition. Column (6) includes all of the controls from columns (2)-(5). All standard errors are clustered at the commander level. $* \, p < 0.1$, $^{**} \, p < 0.05$, $^{***} \, p < 0.01$

Table A9: Impact of a testing frequency by Race and Family Status

	Panel A: Black vs White			
	(1)	(2)	(3)	(4)
	Any Drug Use	Marijuana	Cocaine	Opiates
Testing Frequency	-0.0337	-0.0152	-0.0099	-0.0173
	(0.0309)	(0.0168)	(0.0120)	(0.0143)
Testing Frequency x Black	-0.0825*	-0.0585*	-0.0002	-0.0146
	(0.0433)	(0.0300)	(0.0132)	(0.0155)
Observations (White)	195,141	195,141	195,141	195,141
Dep Var Mean (White)	0.316	0.126	0.059	0.051
Observations (Black)	94,155	94,155	94,155	94,155
Dep Var Mean (Black)	0.562	0.395	0.044	0.054
R^2	0.154	0.154	0.117	0.201
	Panel B: Dependents vs No Dependents			
Testing Frequency	-0.0251	-0.0152	0.0030	-0.0300***
	(0.0274)	(0.0165)	(0.0108)	(0.0107)
Testing Frequency x Has Dependent	-0.0403	-0.0152	-0.0163*	0.0173
	(0.0327)	(0.0211)	(0.0099)	(0.0137)
Observations (No Dependents)	215,684	215,684	215,684	215,684
Dep Var Mean (No Dependents)	0.387	0.222	0.064	0.034
Observations (Has Dependents)	177,021	177,021	177,021	177,021
Dep Var Mean (Has Dependents)	0.329	0.150	0.035	0.071
R^2	0.139	0.138	0.107	0.176

Notes: This table shows the impact of a one-unit increase in commander testing frequency (one additional testing day per month) on soldier outcomes. Panel A compares the impact on black versus white soldiers. Panel B compares soldiers with dependents to soldiers without dependents. Each column includes controls for soldier characteristics as well as fixed effects for the interaction of location, occupation, rank, and arrival quarter and a separate fixed effect for calendar year-month. This is the same specification as table 3 column (3). All standard errors are clustered at the commander level. $* \, p < 0.1$, $^{**} \, p < 0.05$, $^{***} \, p < 0.01$

B Math Appendix

B.1 Proofs

Proposition 1: If being tested is independent of using drugs, the probability that a soldier uses drugs equals the probability that the soldier tests positive conditional on being tested

$$\begin{aligned}\mathbb{E}[TestPos_i|Test_i = 1] &= \mathbb{E}[Test_i \cup DrugUse_i|Test_i = 1]\\ &= \mathbb{E}[DrugUse_i]\end{aligned}$$

The first line follows from the definition of testing positive. The second line follows from conditioning on the soldier being tested.

Corollary 1.1: Regressing testing positive conditional on being tested on testing frequency is the same as regressing Drug use on testing frequency - $CondTestPos_i = \beta^{CTP} CDR_{j(i)} + \mu_i$ $DrugUse_i = \beta^{DU} CDR_{j(i)} + \epsilon_i \implies \beta^{CTP} = \beta^{DU}$

$$\begin{aligned}\beta^{CTP} &= \tfrac{cov(TestPos_i, CDR_{j(i)}|Test_i=1)}{var(CDR_{j(i)})}\\ &= \tfrac{cov(Test_i \cup DrugUse_i, CDR_{j(i)}|Test_i=1)}{var(CDR_{j(i)})}\\ &= \tfrac{cov(DrugUse_i, CDR_{j(i)})}{var(CDR_{j(i)})} = \beta^{DU}\end{aligned}$$

The first line is the definition of β. The second line introduces the definition of testing positive conditional on being tested. The third line simplifies using the fact that we are restricting to soldiers who are tested.

Proposition 2: If being tested is independent of drug use, the probability that a soldier tests positive equals the probability that the soldier uses drugs times the probability that the soldier is tested.

$$\begin{aligned}\mathbb{E}[TestPos_i] &= \mathbb{E}[Test_i \cup DrugUse_i]\\ &= \mathbb{E}[Test_i]\mathbb{E}[DrugUse_i]\end{aligned}$$

Line 1 uses the definition of testing positive. Line 2 uses the fact that being tested is independent of drug use.

Corollary 2.1: The coefficient from regressing testing positive on testing frequency is the same as the coefficient from regressing Drug use on testing frequency multiplied by the probability of being tested - $TestPos_i = \beta^{TP} CDR_{j(i)} + \mu_i$, $DrugUse_i = \beta^{DU} CDR_{j(i)} + \epsilon_i \implies \beta^{TP} = \beta^{DU}\mathbb{E}[Test_i]$

$$\begin{aligned}\beta^{TP} &= \tfrac{cov(TestPos_i, CDR_{j(i)})}{var(CDR_{j(i)})}\\ &= \tfrac{cov(Test_i \cup DrugUse_i, CDR_{j(i)})}{var(CDR_{j(i)})}\end{aligned}$$

$$= \frac{\mathbb{E}[(Test_i \cup DrugUse_i)(CDR_{j(i)})]}{var(CDR_{j(i)})}$$
$$= \frac{\mathbb{E}[(DrugUse_i)(CDR_{j(i)})|Test_i=1]}{var(CDR_{j(i)})} * \mathbb{P}[Test_i = 1]$$
$$= \frac{cov(DrugUse_i, CDR_{j(i)})}{var(CDR_{j(i)})} * \mathbb{P}[Test_i = 1] = \beta^{DU}\mathbb{E}[Test_i]$$

The first line is the definition of a coefficient of a linear regression. The second line introduces the definition of testing positive. The third line introduces the definition of covariance. The fourth line uses the law of iterated expectations. Note that if $Test_i = 0$, $TestPos_i = 0$ so that term disappears. The fifth line uses the fact that being tested is independent of using drugs.

Corollary 2.2: Combining corollary 1.1 and Corollary 2.1 we get that $\beta^{CTP} = \beta^{TP} * \mathbb{E}[Test_i]$.

B.2 Sampling with Replacement

Inspection Tests are conducted under sampling with replacement. That is, a soldier tested on one day can be tested again the following day. Soldiers are removed from the sample following a positive test or a drug or alcohol related criminal offense based on the idea that once a soldier is caught using drugs (or as an alcohol offense) they will be discharged from the Army and no longer have an incentive to not use drugs. This creates a problem as an increase in testing will mechanically catch more drug users independent of any causal response to the increase in drug testing. Therefore, commanders who test more frequently will have fewer drug users in their unit at the end of their command generating a negative mechanical relationship between commander testing frequency and the conditional likelihood of testing positive.

B.2.1 Modeling the Sampling Procedure

The number of drug tests that are positive equals the number of drug users in the unit multiplied by the percent of the unit that is drug tested and the probability of a drug user testing positive when tested.

$$NumPosTest_m = NumDrugUser_m * PercTested * ProbPos$$

The number of drug users in month $m + 1$ equals the number of remaining drug users from month m plus the number of new drug users.

$$NumDrugUser_{m+1} = NewUser + NumDrugUser_m - NumPosTest_m$$

$$= NewUser + NumDrugUser_m - NumDrugUser_m * PercTested * ProbPos$$

$$= NewUser + (1 - PercTested * ProbPos)NumDrugUser_m$$

Now the number of positive tests becomes

$$NumPosTest_m = (NewUser + (1 - PercTested * ProbPos) NumDrugUser_{m-1}) * PercTested * ProbPos$$

Our measure of drug use is the percent of tests that are positive. To find this we need to calculate the number of tests.

$$NumTests = PercTest * NumSoldiers_m$$

The number of soldiers in the unit that are in the sample changes each month based on the number of new soldiers that arrive at the unit and the number of soldiers removed from the sample after a positive test.

$$NumSoldiers_{m+1} = NewSoldiers + NumSoldiers_m - NumPosTest_m$$

$$\implies NumTests_m = PercTest * (NewSoldiers + NumSoldiers_{m-1} - NumPosTest_{m-1})$$

Thus, the percent of positive tests in month m

$$\frac{NumPosTest_{m+1}}{NumTests_{m+1}} = \frac{(NewUser + (1 - PercTested * ProbPos) NumDrugUser_m) * ProbPos}{PercTest * (NewSoldiers + NumSoldiers_m - NumPosTest_m)}$$

If high testing commanders have lower average percent of positive tests, then there will be a negative mechanical correlation between commander testing frequency and the likelihood of testing positive conditional on being tested. We can see in the previous equation that, while an increase in the percent of the unit tested will decrease the numerator, it will have an indeterminate impact on the denominator. In practice, the percent of the unit that tests positive in the following month will decrease.

This can be seen in a simple model of are data. Consider two commanders, A and B. Commander A does the average level of testing while commander B does 12.8% more testing (a 1 unit increase in commander testing frequency increases that likelihood that a soldier is tested by 12.8%). The sample that we use to estimate the impact of testing, consists of soldiers who have newly arrived at each company. Assume that each command lasts 16 months (the median/mean length) and that both companies start with the same number of soldiers, same number of drug users and receive the same number of new soldiers and new drug users each month. Then the difference in the average percent of tests that are positive in each month weighted by the number of tests represents what the coefficient of regressing the likelihood of testing positive on commander testing frequency conditional on being tested if there is no behavioral response to commander testing frequency. We can compare this number to the estimate from Table 3 Column (3) to approximate what percent of the estimate is driven by the sampling with replacement.' Table B1 shows the impact of testing with replacement on the percent of tests that are positive given the above model. Each column varies the percent of drug users that will test positive on a given day and the percent of soldiers that use drugs so that the percent of total tests that are positive is approximately the observed value from inspection tests (0.36-0.37%). If drug users use drugs every day, then all drug users would test positive if tested. However, if a drug user, used cocaine once a month, they would test positive less than 10% of the time tested (cocaine is detectable for 1-3 days after last use, so a soldier who used cocaine once a month would test positive 1-3 days a month). The estimates vary from -0.0147-0.0001 depending on the probability that a drug users tests positive. This is in comparison to our estimate from table 3 of -0.042%. Neither column (1) or column (5) seem entirely reasonable. Column (1) is a world in which very few soldiers ever use drugs, 0.68%, and all drug users are caught the first time they are tested. This is unlikely for multiple reasons. For someone to test positive no matter when they are tested, they would have to be a frequent marijuana user or use other drugs multiple times a week. All soldiers already passed a drug test when they entered the Army. Additionally, we are only using inspection tests. Soldiers that show signs of drug use can be tested under "Probable Cause" which should eliminate positive tests from the heaviest drug users from our sample. Column (5) is unlikely for a different reason. If drug users only test positive 1% of the time tested (this would occur if drug users use drugs once per year), over 38% of soldiers would need to use drugs in order to have 0.36% of tests come back positive. Columns (3) and (4) seem the most likely. Here we are assuming that about 2-4% of soldiers use drugs and that drug users use drugs once or twice a month. This still suggests that we are overestimating the impact of commander testing frequency on drug use by 4-10%.

Table B1: Impact of commander testing frequency on observed drug use caused by sampling with replacement

	(1)	(2)	(3)	(4)	(5)
Mechanical Bias	-0.0147	-0.0084	-0.0043	-0.0017	-0.0001
Percent of Tests Positive	0.378	0.375	0.368	0.361	0.355
% Drug Users	0.68	1.06	1.82	4.10	38.30
Probability of Testing Positive	100%	50%	25%	10%	1%

Notes: This table shows change in the likelihood of a soldier testing positive conditional on being drug tested caused by a 12.8% increase in the monthly testing and sampling with replacement. 12.8% represents the increase in the likelihood of being tested caused by a one standard deviation increase in commander testing frequency. This can be compared to the coefficient from table 3 column (3) to estimate what percent of our estimate of the impact of commander testing frequency on drug use is driven by sampling with replacement and not behavior responses by soldiers to the increased likelihood of being tested. Each column varies the percent of soldiers that are drug users and the probability that a drug user will test positive on a given day so that a similar percent of tests will be positive as observed in the data.

B.2.2 Simulating the Mechanical Bias

We can also estimate the mechanical bias use by simulating the testing environment. Under the same assumptions as in the previous exercise (the average commander tests 17.34% of soldiers each month, the average command lasts 16 months and about 0.36-.37% of tests are positive), we simulate 1 million soldiers who are either assigned to an average testing commander or a high testing commander. We randomly assign soldiers to each command uniformly across different months of the command. Being assigned to the high testing commander increases the likelihood of being tested by 12.8%. We randomly identify a percentage of soldiers as drug users. We randomly select soldiers to be tested with the likelihood that each soldier is tested corresponding to their commanders testing propensity. We vary the likelihood that a drug user will test positive if tested each month to correspond to different types of drug use. We report Bootstrap 95% confidence intervals. While the confidence intervals are somewhat wide, the point estimates are similar the estimates of mechanical bias from our model in table B1. Again, our preferred specifications are in columns (3) and (4) which assume that 2-4% of soldiers are using drugs once or twice a month show that we are overestimating the impact of drug testing on illicit drug use by 3-10%.

Table B2: Impact of commander testing frequency on observed drug use caused by sampling with replacement

	(1)	(2)	(3)	(4)
Mechanical Bias	-0.0156	-0.0089	-0.0044	-0.0013
Bootstrap 95% CI	[-0.0222,-0.0094]	[-0.0152,-0.0021]	[-0.0105,0.0020]	[-0.0074,0.0049]
Mean	0.388	0.386	0.377	0.368
Observations	10,000,000			
Iterations	500			

Notes: This table simulates change in the likelihood of a soldier testing positive conditional on having a drug test caused by a 12.8% increase in the monthly testing and sampling with replacement. 12.8% represents the increase in the likelihood of being tested caused by a one standard deviation increase in commander testing frequency. This can be compared to the coefficient from table 3 column (3) to estimate what percent of our estimate of the impact of commander testing frequency on drug use is driven by sampling with replacement and not behavior responses by soldiers to the increased likelihood of being tested. Each column varies the percent of soldiers that are drug users and the probability that a drug user will test positive on a given day so that a similar percent of tests will be positive as observed in the data.