Managerial Discretion and Racial Disparities: Evidence from Military Drug Screening

Kyle Greenberg[1] Carson Homme[1]
Anika Nerella[2] **CarlyWill Sloan**[1]

The views expressed in this presentation are those of the authors and do not necessarily represent the views, official policy, or position of the United States Military Academy, the Department of the Army, the Department of Defense, or the U.S. Government. Please do not cite or distribute without the authors' permission.

[1]USMA at West Point

[2]University of Pennsylvania

Request

Please don't share on social media.

Big Picture: Disparities & Discrimination

Racial disparities are readily apparent across many settings in the US (e.g., labor markets, the criminal justice system, healthcare).

While estimating racial disparities is straightforward, identifying discrimination is more challenging.

- This is because researchers rarely can compare individuals of different races who have identical baseline behaviors (here actual drug use rates).

In this paper, we will ask whether racial disparities in managerial drug testing persist *after accounting for drug usage*

Why Focus on Manager Drug Testing?

High-Stakes Managerial Decision:

- One of the most consequential discretionary actions managers can take
- $\rightarrow$ Failure may result in separation or job loss

Widespread Civilian Use:

- 41% of U.S. employers have punitive drug testing policies (NSDUH)
- $\rightarrow$ Especially common in public safety and security sectors

Policy Implications:

- Racial disparities in discretionary enforcement may affect retention, discipline, and mobility
- $\rightarrow$ Relevant for both military careers and civilian transitions

Where we fit: Research questions

Aim to answer three questions

1. Are there *average* racial disparities in drug testing?
2. Are there *average* racial disparities in drug testing **after controlling for drug usage**? $\rightarrow$ discrimination
3. Is there significant heterogeneity in discrimination across brigades?

What we do: Our Setting

We observe the following:

1. **Drug usage** through random drug tests & failures
   - Army's mandatory and extensive *random* drug testing program allows commanders to see "true" drug usage by subgroup
     - We show empirical evidence support for randomness in the paper
   - Measured at the battalion level (about 750 soldiers)
2. **Commander decision making** through "discretionary" drug tests
   - Uniformed Code of Military Justice affords commanders a legal right to drug test at any time with probable cause
   - Commanders are only correct 25% of the time

Who are Commanders?

Company Commander

Battalion Commander

Commanders

Commanders (company and battalion) are the Army's managers

Aspect	Role	Military Commander	Civilian Equivalent
Battalion	Demographics	>90% White, 41, Lt. Col	-
	Oversight	about 750 soldiers	Division Head (David Wallace)
	Drug Testing	Sets policy, reviews results	Sets compliance, reviews metrics
	Admin Power	Can initiate separation	Can initiate termination
Company	Demographics	>90% White, 31, Captain	-
	Oversight	about 100 soldiers	Supervisor (Michael Scott)
	Drug Testing	Implements testing	Implements protocols
	Admin Power	Recommends discipline	Recommends HR action

$\rightarrow$ **Both can order discretionary tests**

Where we fit: Other work

Research on managerial discretion shows discrimination affects hiring/performance/recognition, but less is known about daily decisions affecting worker discipline, especially in public sector settings.

e.g., Bertrand and Schoar (QJE 2003); Bloom and Van Reenen (QJE 2007); Bloom et al. (QJE 2011); Bandiera et al. (JPE 2020); Lazear, Shaw, and Stanton (JOLE 2015); Baltrunaite, Bovini, and Mocetti (WP 2020); Fenizia (Econometrica 2022), Minni (WP 2023).
Glover, Pallais, and Pariente (QJE 2017); Benson, Board, and Meyer-ter-Vehn (REStud 2024); Rim, Rivera, Kiss, and Ba (JOLE 2024)

Where we fit: Other work

Research on managerial discretion shows discrimination affects hiring/performance/recognition, but less is known about daily decisions affecting worker discipline, especially in public sector settings.

e.g., Bertrand and Schoar (QJE 2003); Bloom and Van Reenen (QJE 2007); Bloom et al. (QJE 2011); Bandiera et al. (JPE 2020); Lazear, Shaw, and Stanton (JOLE 2015); Baltrunaite, Bovini, and Mocetti (WP 2020); Fenizia (Econometrica 2022), Minni (WP 2023).

Glover, Pallais, and Pariente (QJE 2017); Benson, Board, and Meyer-ter-Vehn (REStud 2024); Rim, Rivera, Kiss, and Ba (JOLE, 2024)

e.g., Arnold, Dobbie, and Hull (QJE 2022); Canay, Mogstad and Mountjoy (ReStud 2024); Goncalves and Mello (AER 2021); Luh (WP 2023); Ridgeway and MacDonald (JASA 2009); Lang and Kahn-Lang Spitzer (JEP 2020)

Where we fit: Other work

We think this paper bridges work on managerial discretion and racial bias in rule-based institutions, showing that enforcement disparities are widespread—nearly always disadvantaging Black soldiers—but vary substantially in magnitude across managers.

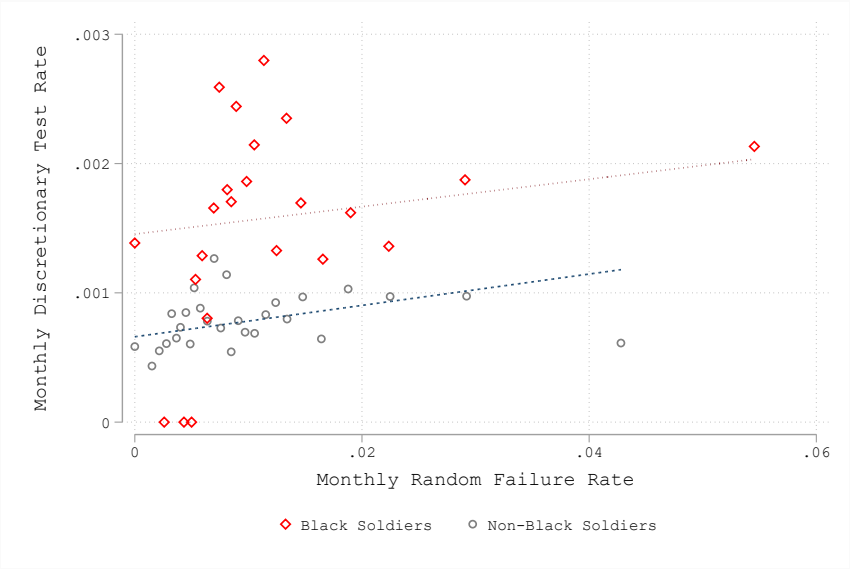
Data & Sample Construction

- **Panel:** Monthly panel of all enlisted Army soldiers (2015–2022), linked to drug test and administrative records.
- **Unit of Observation:** Soldier-month
- **Key Linkages:**
  - Drug test results (type, date, outcome)
  - Personnel records: race, AFQT score, occupation, unit history
  - Each soldier-month linked to a battalion-commander spell (avg. duration $\approx$ 2 years)
- **Sample Restrictions:**
  - First-term, unmarried, male junior enlisted soldiers (likely to reside in barracks)
  - Drop soldier-months after drug offenses
  - Drop training/support units and commanders with $<$12 months tenure
  - End with $N = 2.53$ *million*

Summary Stats

	All (N=172,339)	Black Soldiers (N=29,902)	Non-Black Soldiers (N=142,437)
Counts			
Discretionary Tests	2,469	697	1,772
Discretionary Test Fails	582	209	373
Random Tests	416,792	64,719	352,073
Random Test Fails	2,897	835	2,062
Random Failure Rate Q1	0%	0%	0%
Discretionary Test Rate Q1	0.087%	0.142%	0.061%
Random Failure Rate Q2	0.271%	0.367%	0.270%
Discretionary Test Rate Q2	0.062%	0.122%	0.062%
Random Failure Rate Q3	0.539%	0.554%	0.538%
Discretionary Test Rate Q3	0.096%	0.171%	0.093%
Random Failure Rate Q4	0.860%	0.903%	0.856%
Discretionary Test Rate Q4	0.112%	0.196%	0.103%
Random Failure Rate Q5	1.966%	2.643%	1.583%
Discretionary Test Rate Q5	0.130%	0.189%	0.096%

Discretionary Tests and Failure Rates by Race



Regression

We estimate our results in a solider-month panel, clustering at the battalion level

$$DiscretionaryTest_{it} = \alpha + \beta BlackSoldier_i + \gamma D_{br(it)} + \delta R_{brk(it)} + X_i'\Gamma + \pi_t + \epsilon_{it} \quad (1)$$

To answer the question:
Do racial disparities exist ($\beta ! = 0$)? after controlling for **drug usage by race**

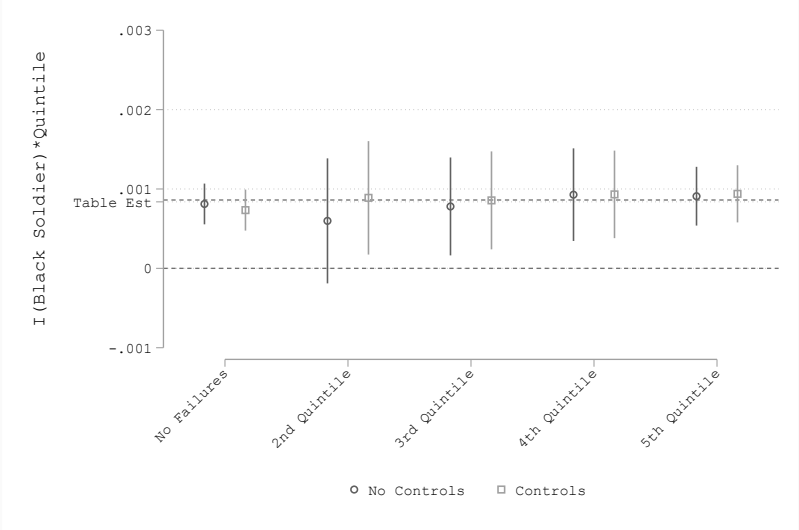
- i.e., random drug test failure rates by race
- Random test failure rates at the battalion race level ($D_{br(it)}$)
- Random test failure rates at the battalion race-occupation-entry year level ($R_{brk(it)}$)

Regression Results

	Outcome: Discretionary Test			
	(1)	(2)	(3)	(4)
Black Soldier=1	0.0947***	0.0861***	0.0861***	0.0814***
	(0.00760)	(0.00790)	(0.00791)	(0.00797)
Battalion x Race Random Fail Rate		1.125***	0.865**	0.840**
		(0.374)	(0.378)	(0.376)
Battalion x Race x MOS x Entry Yr			0.261***	0.258***
Random Fail Rate			(0.0841)	(0.0841)
N	2530976	2530976	2530976	2530976
NonBlack Soldier Mean*100	0.083	0.083	0.083	0.083
% NonBlack Soldier Mean	114	103	103	98
(MOS)x(Entry Yr) FE, Month FE	Y	Y	Y	Y
Soldier Controls				Y

→Black soldiers face twice as much testing, despite controlling for drug use patterns.

Results by Drug Usage



→ Over-testing of Black soldiers is similar across battalion drug usage rates.

Measurement Error (?)

In the paper we:

- Restrict to larger battalions (>3,050 soldier-months)
- Examine battalions with balanced Black/non-Black populations to control for sample size differences
- Analyze battalions with similar drug failure rates between racial groups

$\rightarrow$ Results are robust to each sample restriction

Further, random testing is widespread (over 400,000 tests in our sample) and most drug charges come from drug test failures

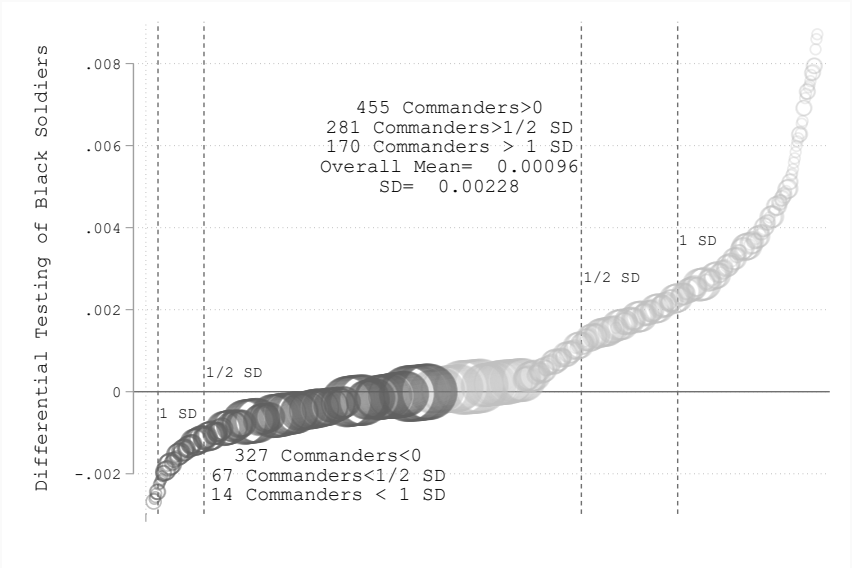
Is this widespread?

Are Disparities Widespread or Driven by a Few?

1. **Estimate testing disparities for each battalion commander**
   – Measure how often each commander over-tests Black vs. non-Black soldiers
2. **Aggregate results to the brigade level**
   – Summarize patterns across units to see where disparities are concentrated
   (669 battalions $\rightarrow$ 33 brigades)
3. **Use Empirical Bayes to adjust for noise**
   – Shrinks noisy estimates toward the mean, revealing true underlying variation

$\rightarrow$ *Helps us identify whether racial disparities are systemic or concentrated among a few commanders.*

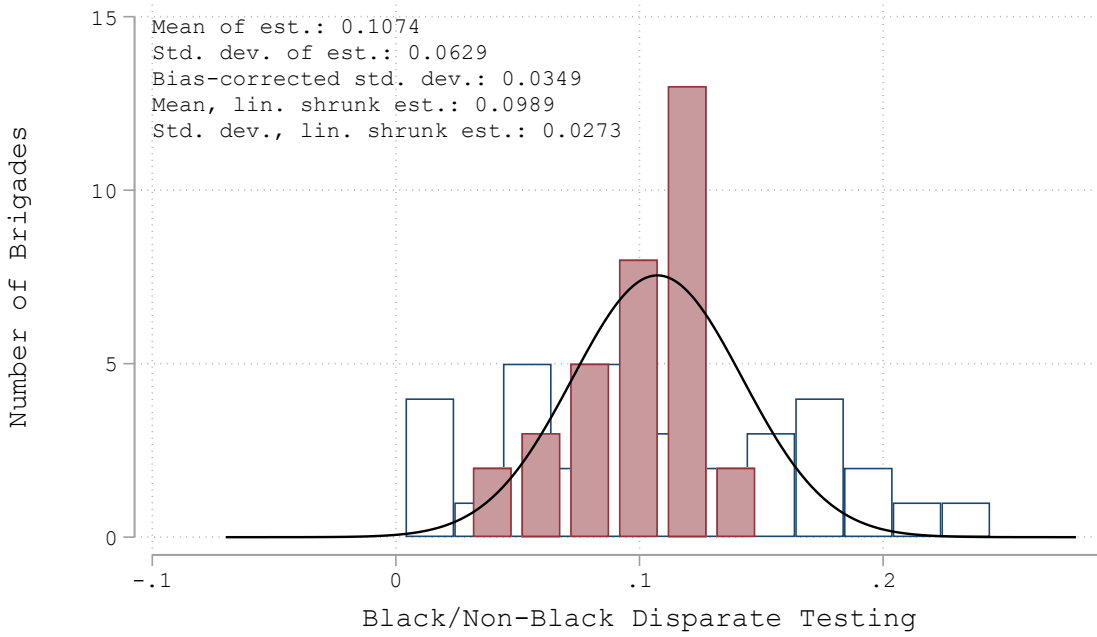
Battalion Commander-Level Estimates: Overall

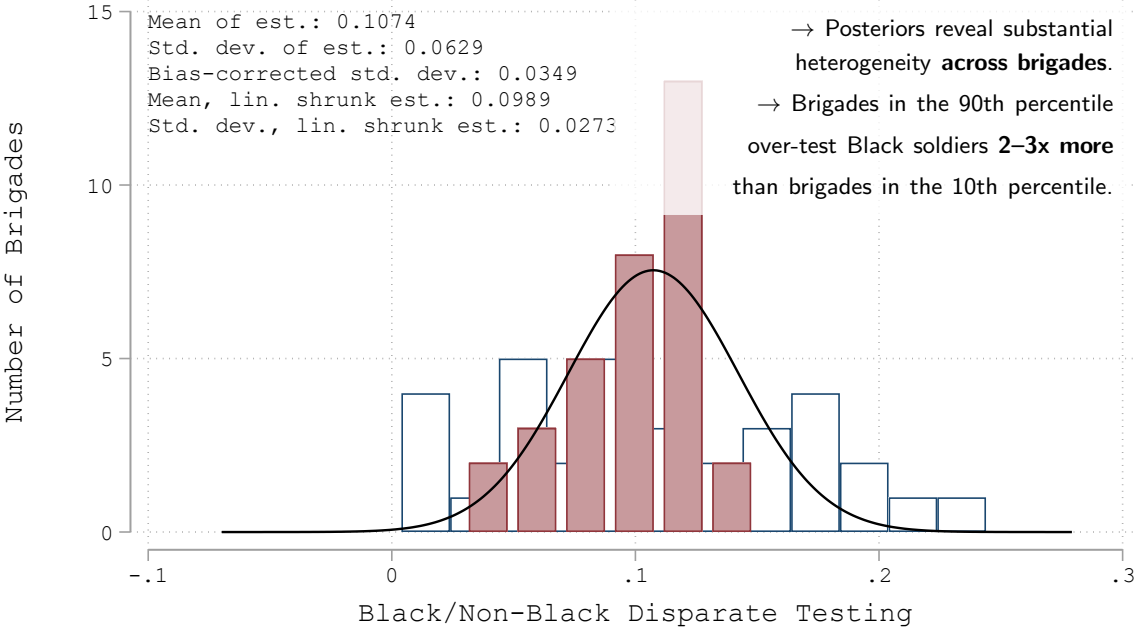


Empirical Bayes Estimation of Disparate Testing

We apply Empirical Bayes methods to adjust for statistical noise and recover the true underlying disparities following *Kline et al. 2022; Kline et al. 2024; Walters 2024*.

1. Average commander-level estimates within brigades, weighted by soldier-months.
2. Estimate the variance of true disparities across brigades, correcting for bias.
3. Apply **linear shrinkage**—a weighted average of each brigade's raw estimate and the overall mean
4. Adjust for **multiple testing** by estimating the share of true nulls and computing **q-values**. $\rightarrow$ 22/33 brigades exhibit statistically significant evidence of disparate testing


Mean of est.: 0.1074
Std. dev. of est.: 0.0629
Bias-corrected std. dev.: 0.0349
Mean, lin. shrunk est.: 0.0989
Std. dev., lin. shrunk est.: 0.0273
Number of Brigades
15
10
5
0
-.1
0
.1
.2
Black/Non-Black Disparate Testing


Mean of est.: 0.1074
Std. dev. of est.: 0.0629
Bias-corrected std. dev.: 0.0349
Mean, lin. shrunk est.: 0.0989
Std. dev., lin. shrunk est.: 0.0273
→ Posteriors reveal substantial heterogeneity across brigades.
→ Brigades in the 90th percentile over-test Black soldiers 2–3x more than brigades in the 10th percentile.
Number of Brigades
15
10
5
0
-.1
0
.1
.2
.3
Black/Non-Black Disparate Testing

The End

Main Findings:

- Black soldiers are **twice as likely** to face discretionary drug testing compared to non-Black soldiers, despite accounting for drug usage rates.
- Disparities vary across brigades, but **the majority over-test** Black soldiers.

Why We Should Care:

- **Economic Mobility**: discrimination in testing impacts Black soldiers' careers
- **Public Sector Relevance**: Military veterans, who make up a significant portion of police officers, may bring practices learned in the military that contribute to racial disparities and inequalities in civilian law enforcement and justice outcomes.
- **Broader Labor Market**: Equitable implementation of workplace policies is crucial as drug testing is common.

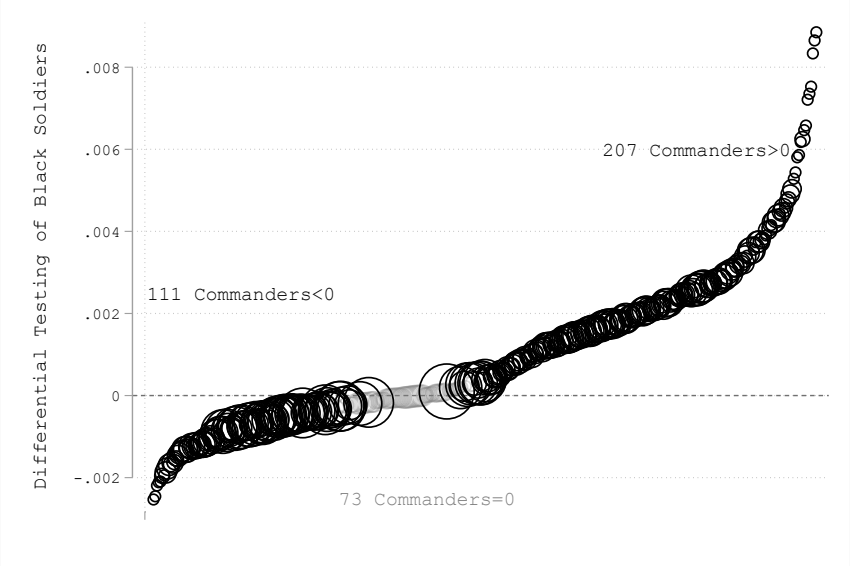
Frequently Asked Questions

- What types of drugs are driving your results? **Marijuana**.
- Can the disparity in drug usage be explained by the amount of time a drug stays in your system? **No**. Mayo Clinic Proceedings (2008)
  - Marijuana (3 days for single use)
  - Cocaine metabolites (2–4 days)

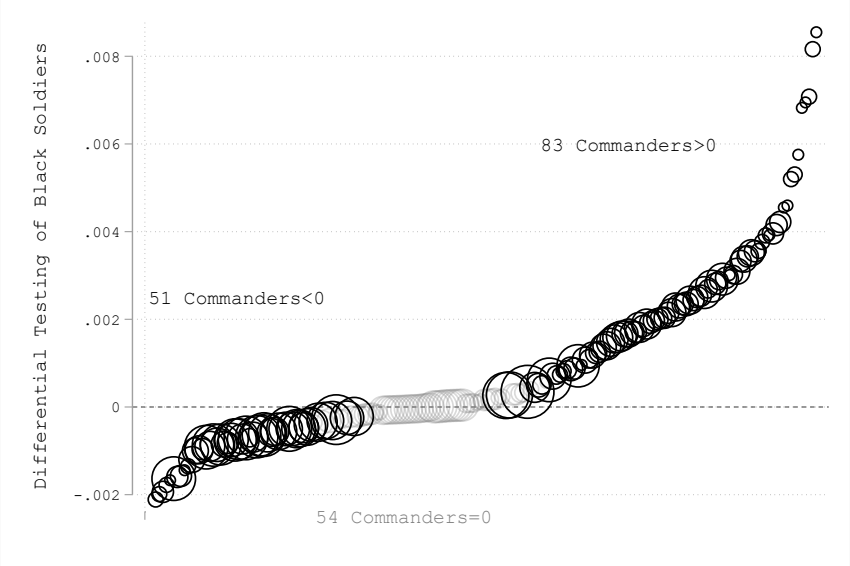
TABLE 2. **Length of Time Drugs of Abuse Can Be Detected in Urine**

Drug	Time
Alcohol	7-12 h
Amphetamine	48 h
Methamphetamine	48 h
Barbiturate	
Short-acting (eg, pentobarbital)	24 h
Long-acting (eg, phenobarbital)	3 wk
Benzodiazepine	
Short-acting (eg, lorazepam)	3 d
Long-acting (eg, diazepam)	30 d
Cocaine metabolites	2-4 d
Marijuana	
Single use	3 d

Larger Battalions



Battalions with similar Black and non-Black failure rates



References

Kline, Patrick, Evan K Rose, and Christopher R Walters (2022). "Systemic discrimination among large US employers". In: *The Quarterly Journal of Economics* 137.4, pp. 1963–2036.

— (2024). "A Discrimination Report Card". In: *American Economic Review* 114.8, pp. 2472–2525.

Walters, Christopher (2024). "Empirical Bayes methods in labor economics". In: *Handbook of Labor Economics*. Vol. 5. Elsevier, pp. 183–260.