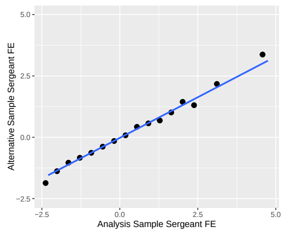
Supplementary Appendix for The Boss in Blue: Supervisors and Police Behavior

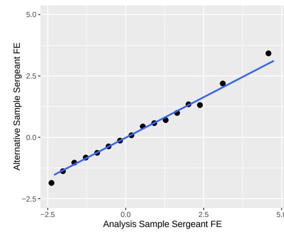
A Additional Figures

Figure A.1: Robustness to Alternative Sampling Decisions

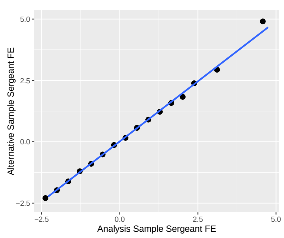
(a) Unrestricted
Correlation = 0.8720



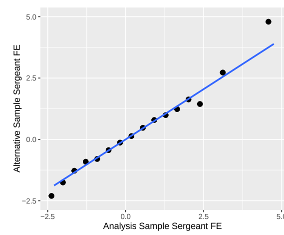
(b) Impute Missing Within Spell
Correlation = 0.8666



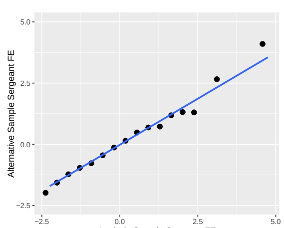
(c) Impute Missing Within Spell, remove everything else
Correlation = 0.9873



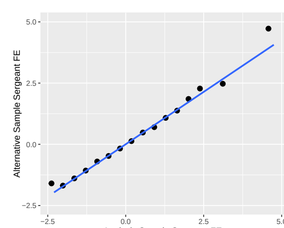
(d) Impute all temporary assignments
Correlation = 0.9362



(e) Keep temporary, remove missing
Correlation = 0.9153

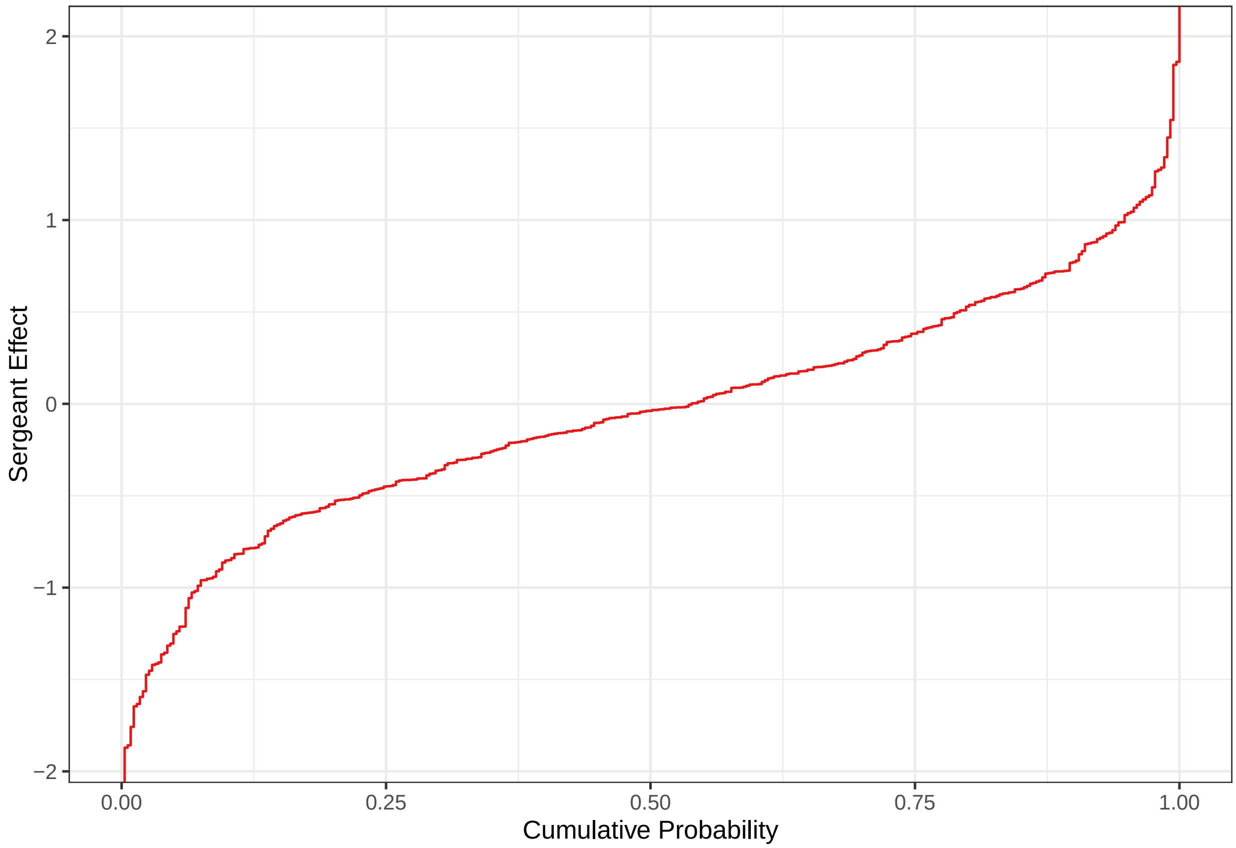


(f) Keep missing, remove temporary
Correlation = 0.9311



Notes: This figure presents the correlation between sergeant fixed effects under different sampling restrictions. (a) makes no sample restrictions, (b) imputes missing months within a sergeant spell and keeps any other observations where the sergeant is unknown, (c) is the same as (b) but other observations with unknown sergeants are removed, (d) imputes temporary one-off assignments with different sergeants using an officer's permanent sergeant, (e) keeps all of the temporary (single month) assignments but removes all months with an unknown sergeant, and (f) keeps months with an unknown sergeant but removes the temporary (single month) assignments. In the specifications that include months with an unknown sergeant, the model includes an 'Unknown Sergeant' fixed effect.

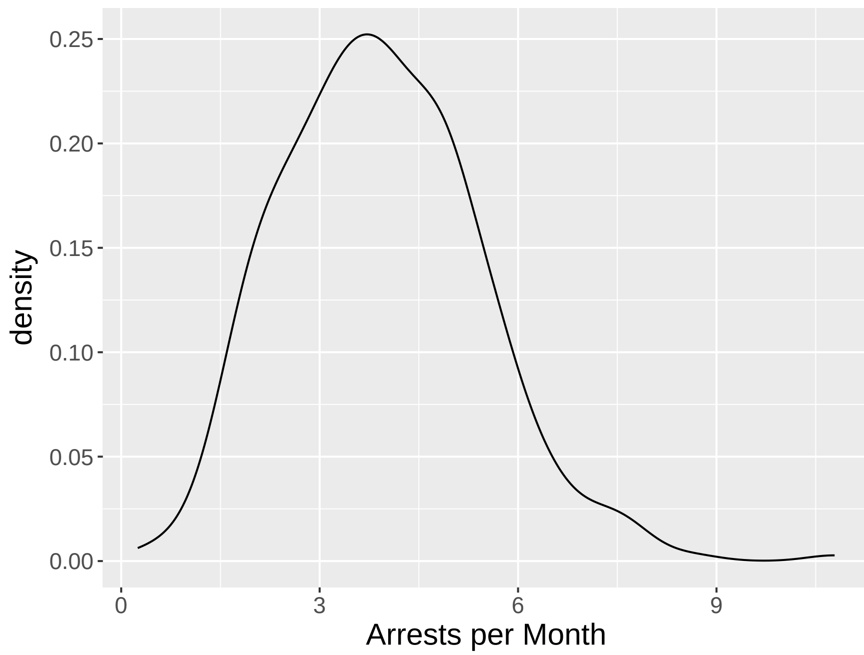
Figure A.2: CDF of Supervisor Fixed Effects



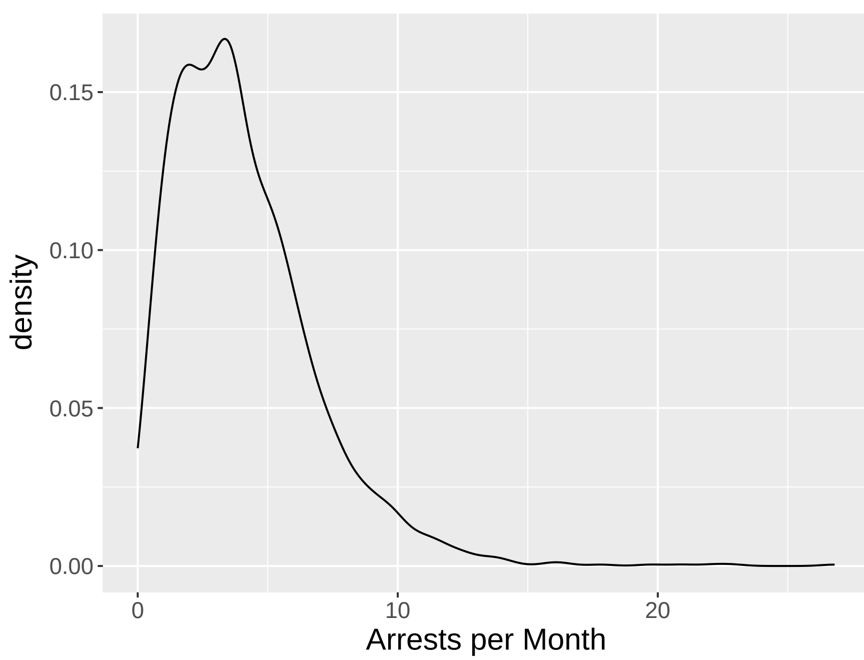
Notes: This figure displays the CDF of the sergeant effects, estimated using the sergeant fixed effects in equation 1 that are multipled by the Bayesian shrinkage factor described in Section 4.

Figure A.3: Distribution of Arrests

(a) Across Sergeants

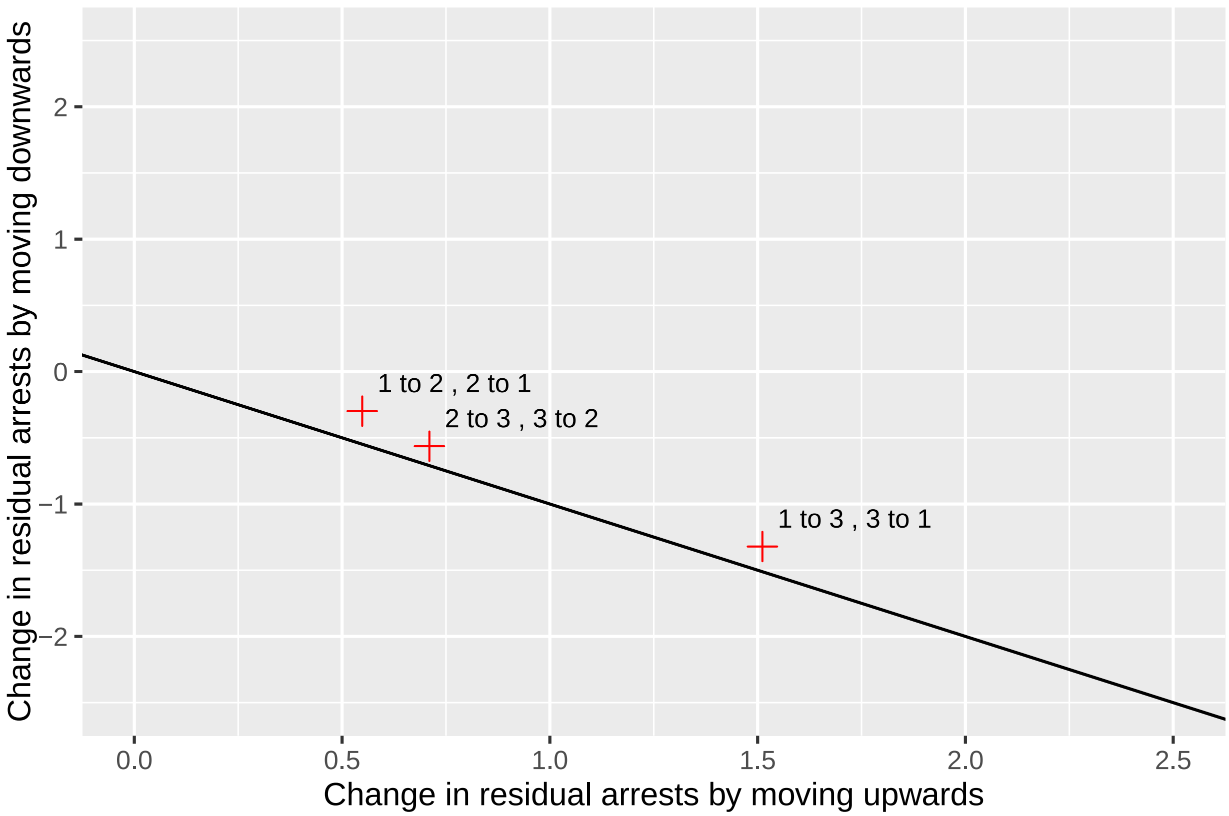


(b) Across Officers



Notes: These figures present empirical distributions for monthly arrests. For a given sergeant, I calculate the average number of monthly arrests made by officers who work for them. A.3a then plots the distribution of this average. Then, for each officer, I calculate the average number of arrests they make in a month across all months they are in the sample. I plot the distribution of this average in A.3b.

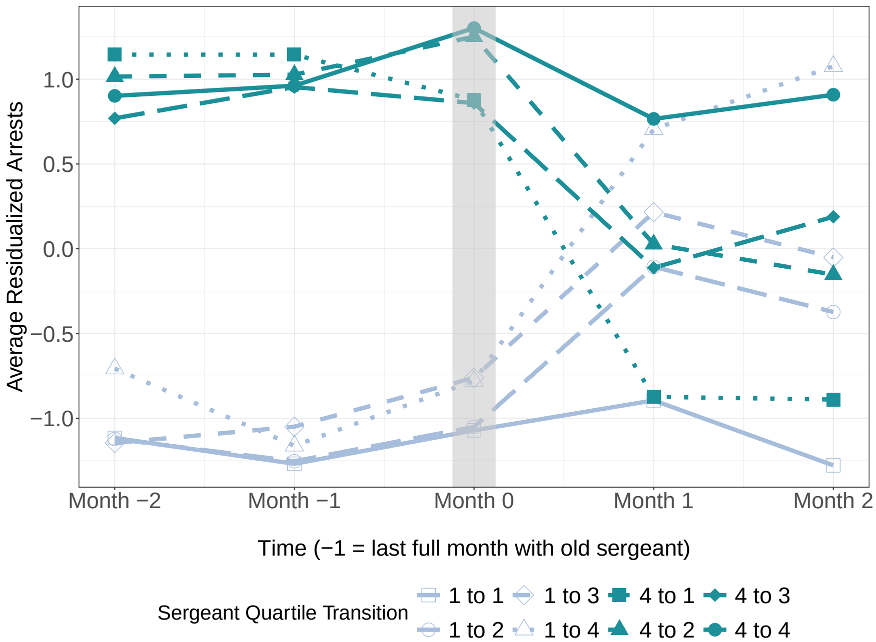
Figure A.4: Symmetry in moves



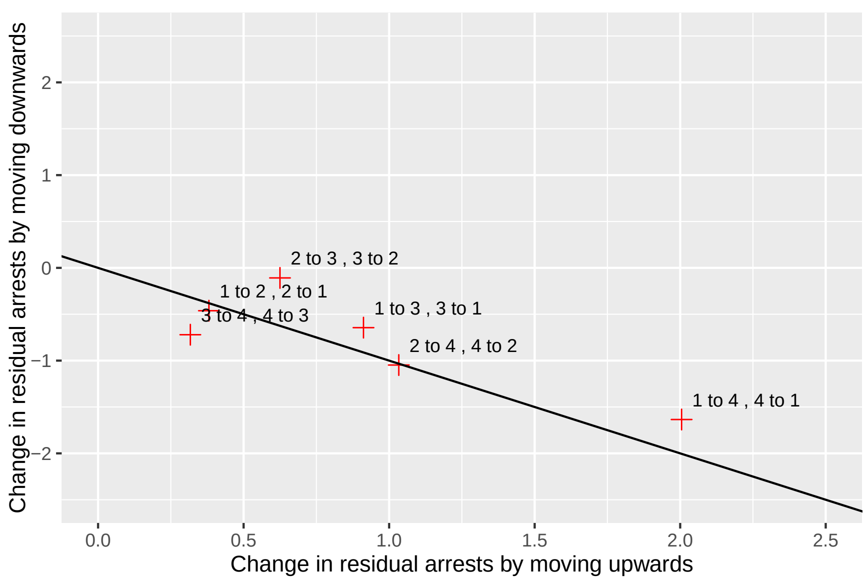
Notes: Each crosshair represents a pair of symmetric moves between sergeants in different terciles of average residualized arrests. Changes in residual arrests are calculated as the average difference between the average number of arrests 2 months after a move and 2 months before a move. Arrests are residualized by officer, sector-watch, and day-off group fixed effects and a second degree polynomial of tenure, as described in Section 4.

Figure A.5: Event study using sergeant quartiles

(a) Event Study Figure

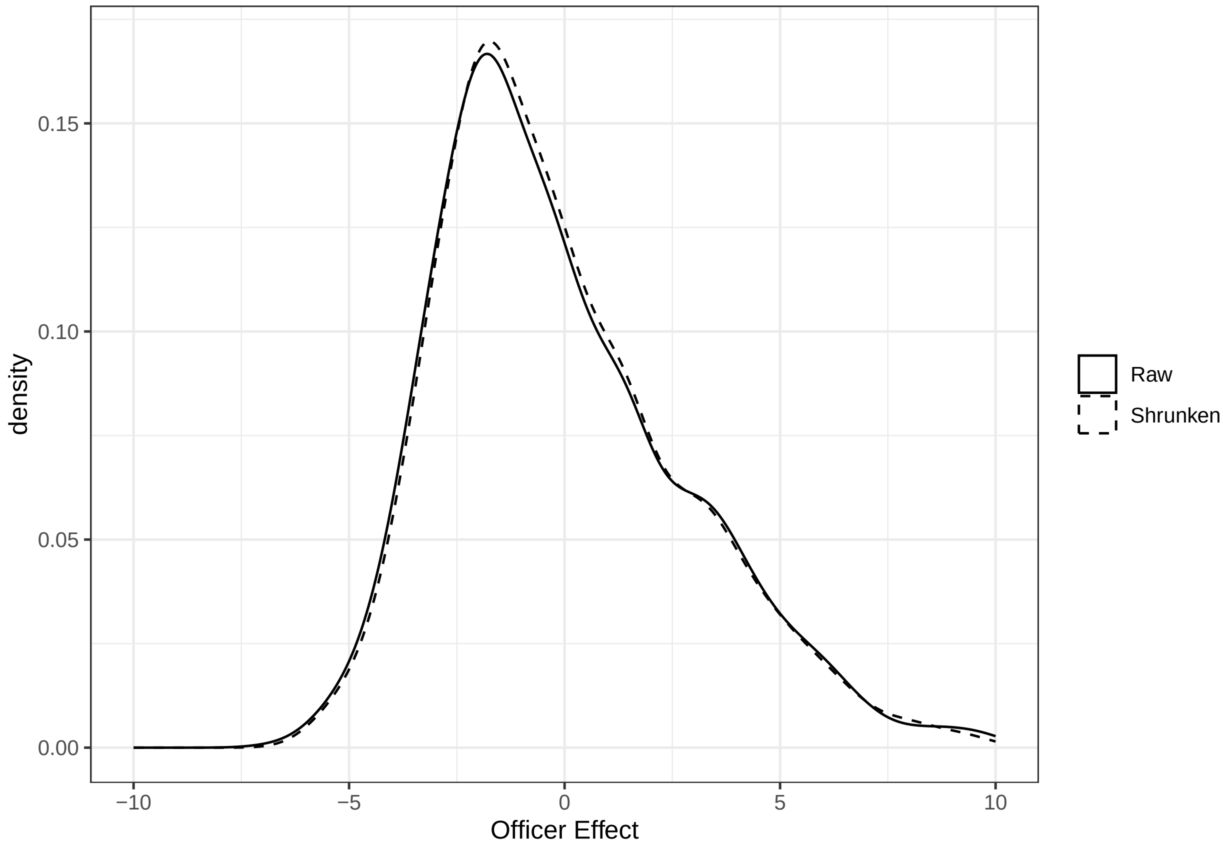


(b) Symmetry Test



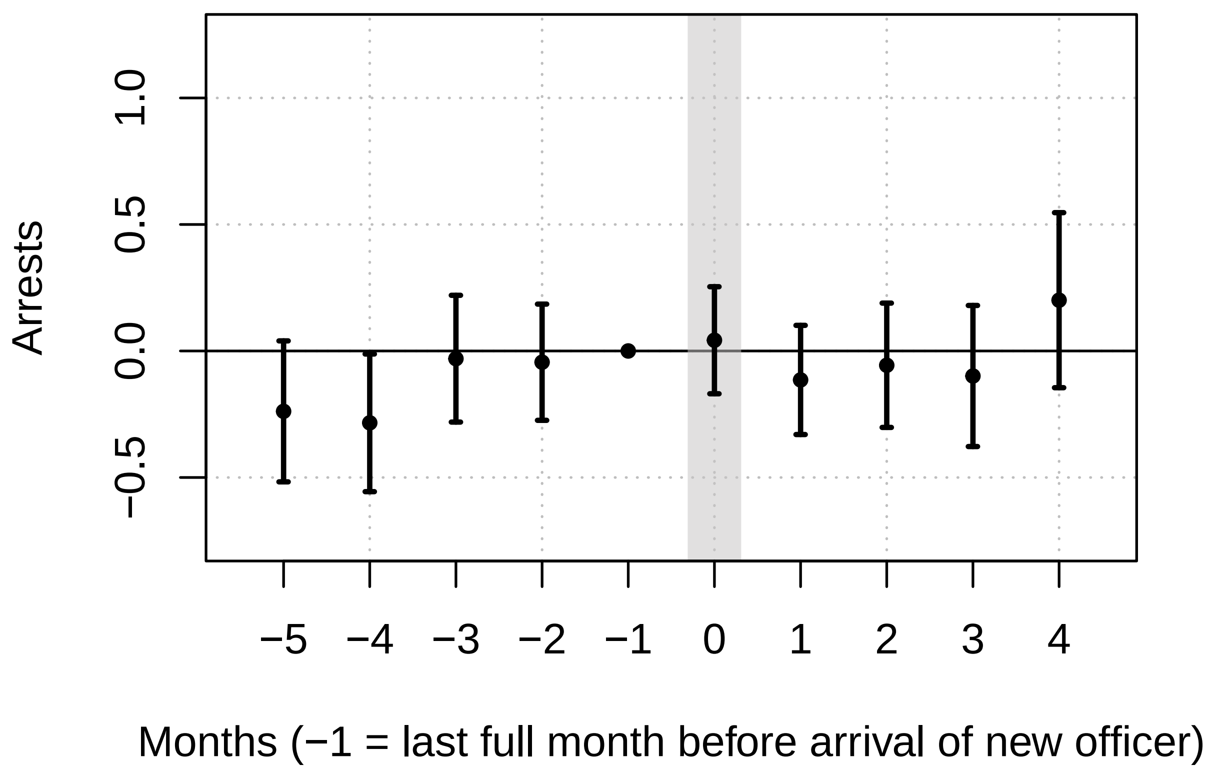
Notes: These figures present the same information as Figures 1 and A.4, instead splitting supervisors into quartiles rather than terciles. To limit the amount of lines in A.5a, I only plot transitions from supervisors in the highest and lowest quartiles. The symmetry test uses all symmetric quartile transitions. In A.5b, the crosshairs align roughly with the -45 degree line, suggesting the presence of symmetry across moves in equal and opposite directions.

Figure A.6: Distribution of Officer Effects



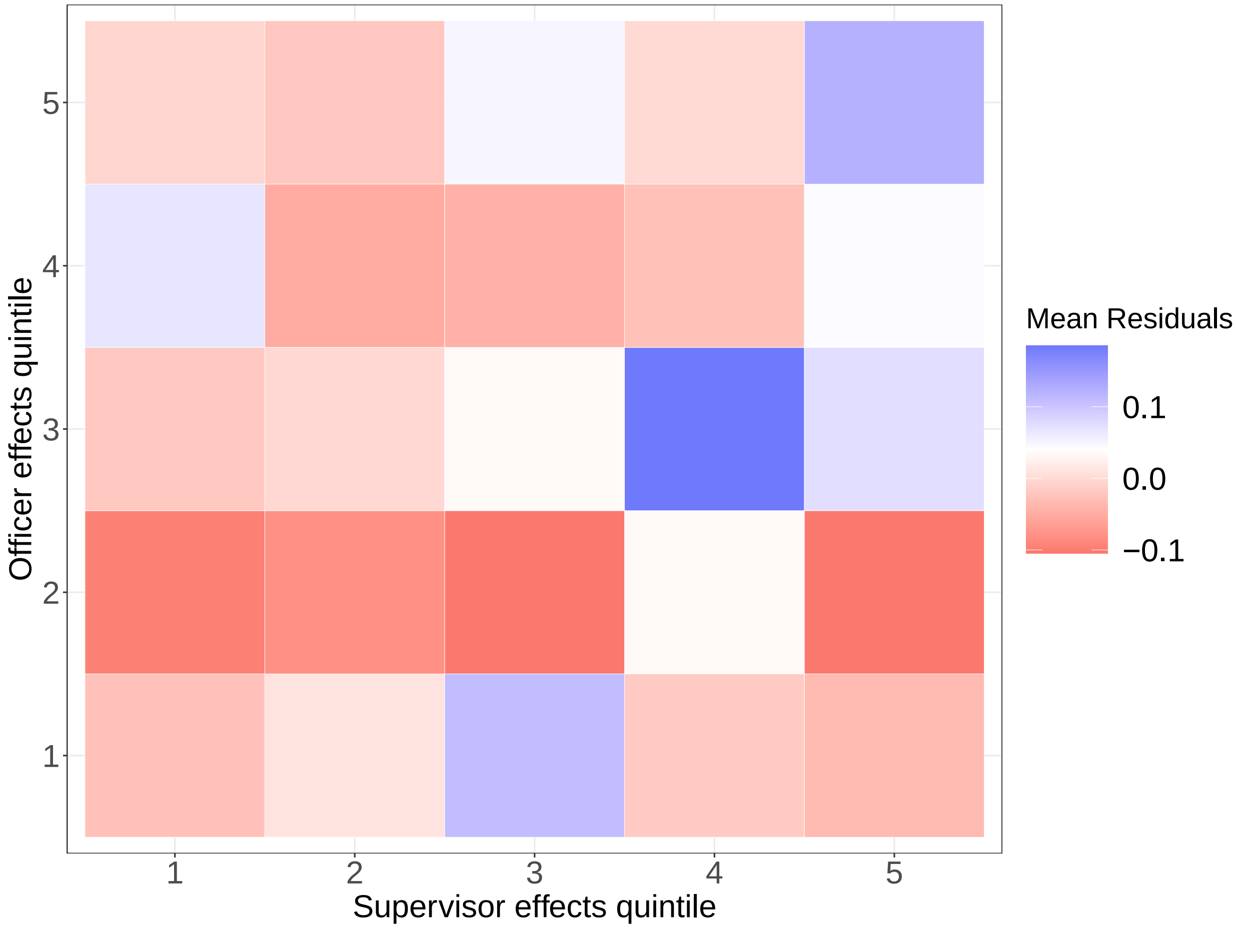
Notes: This figure plots the officer fixed effects estimated using equation 1. The solid line presents the raw fixed effects, while the dotted line presents the raw effects multiplied by the Bayesian shrinkage factor as described in Section 4.

Figure A.7: Arrests Made by Incumbent Officers



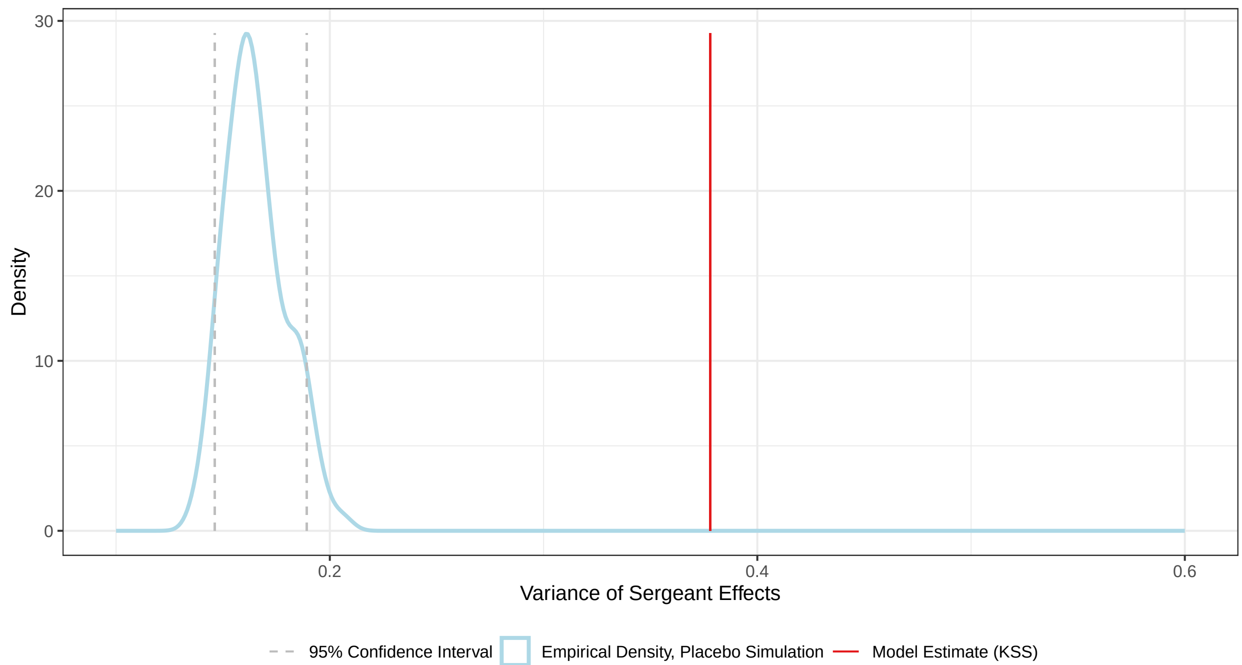
Notes: This figure plots the event study coefficients from equation 5. For an officer switching event e in which officer i switches from sergeant j to sergeant j', incumbent officers are those who work with sergeant j for 5 months before the event and and 4 months after the event. The x-axis indicates months relative to officer i's switch. Standard errors are clustered at the level of the switching officer.

Figure A.8: Residuals by quintile of officer and sergeant arrest effects



Notes: This figure reports the average residuals by quintiles of officer and sergeant (supervisor) arrest effects. Darker blue indicates more positive residuals and darker red indicates more negative residuals.

Figure A.9: Placebo Test, Sergeants randomly reassigned to officers



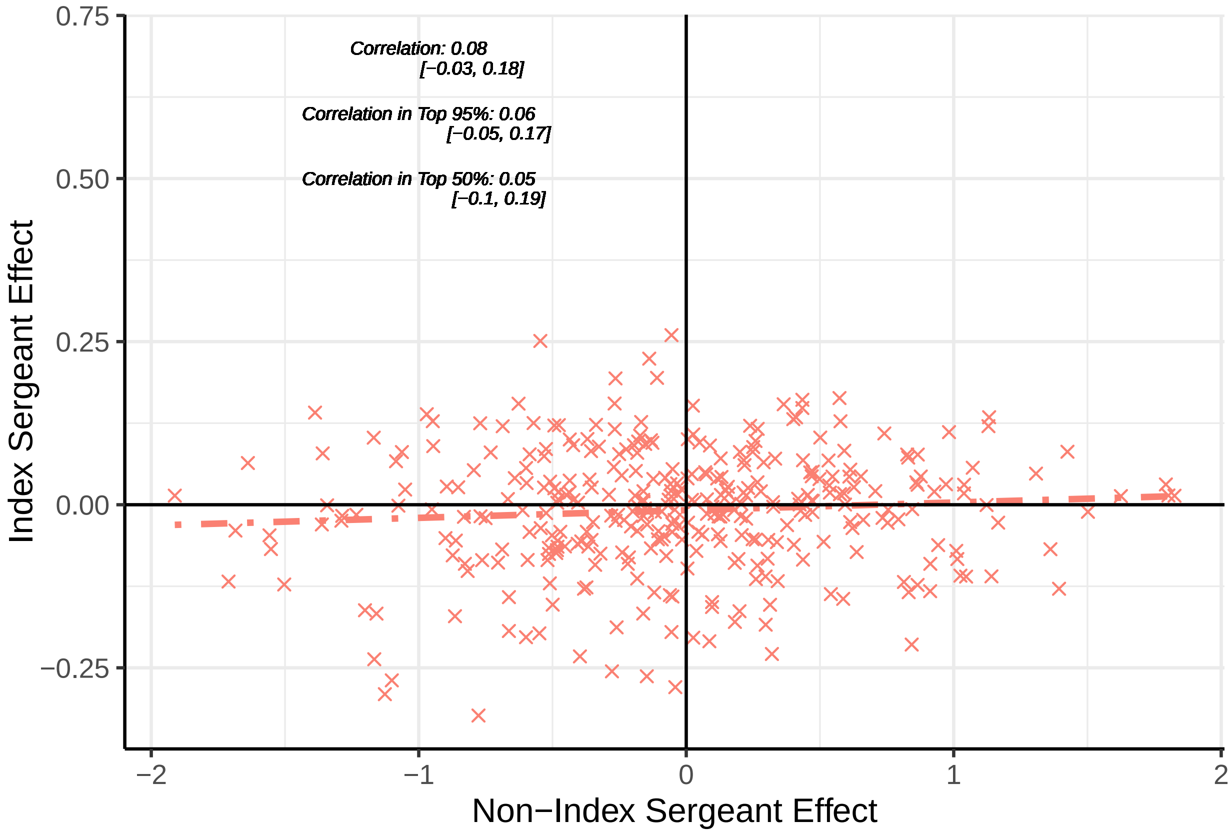
Notes: This figure depicts the results of placebo tests that randomly reallocate sergeants to officers, preserving the number of unique officers managed for each sergeant. For every reallocation, I estimate equation 1 and report the resulting (unadjusted) variance of sergeant effects. The empirical density, in light blue, plots the density of variance estimates for 100 reallocations. The dashed lines denote the 95% confidence interval of the placebo variance estimates. The red vertical line denotes my main estimate of the variance in sergeant effects, adjusted for bias using the KSS method described in Section 4.

Figure A.10: Distribution of Sergeants between Serious and Low-level Effects

Tercile of Serious Sergeant Effect			
3	11.2%	11.2%	11.0%
2	7.8%	12.7%	12.7%
1	14.4%	9.2%	9.8%
	1	2	3
	Tercile of Low-level Sergeant Effect		

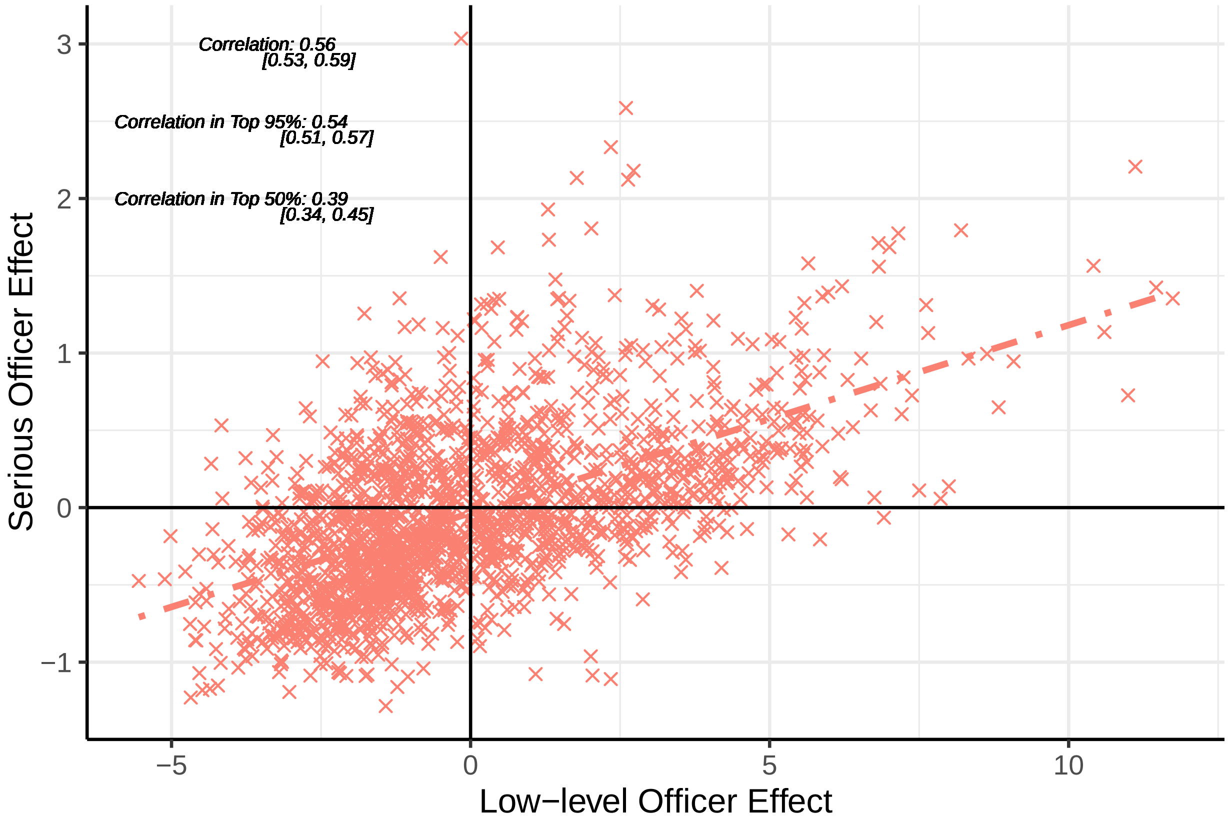
Notes: This figure displays the percentage of sergeants within each tercile grouping of (Bayes-shrunken) low-level and serious sergeant effects.

Figure A.11: Relationship between non-index and index sergeant effects



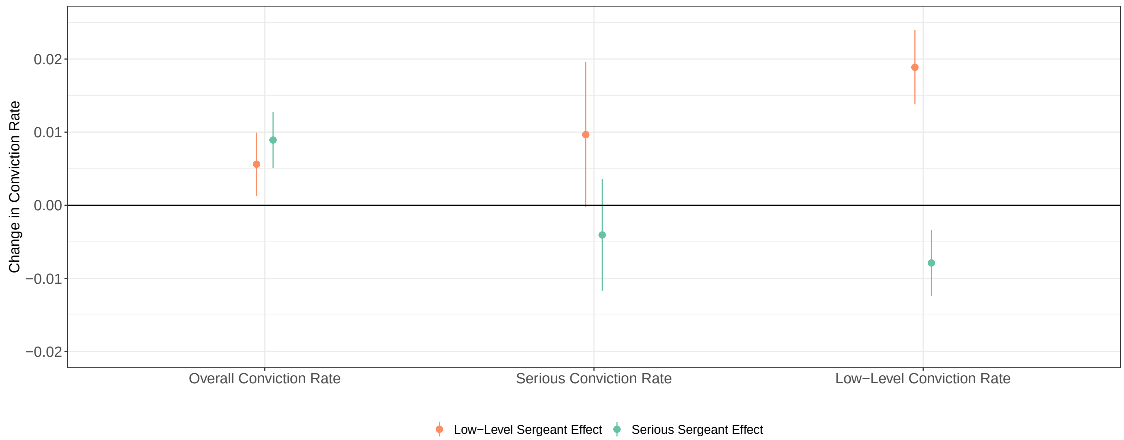
Notes: This figure displays a scatterplot of the relationship between index sergeant effects and non-index sergeant effects, each representing a sergeant's propensity to induce arrests for crimes of either type, estimated using equation the two-way fixed effects model in equation 1. These measures are used as a robustness check on serious and low-level arrest classifications used to generate Figure 5. Each point represents a sergeant. Index crimes encompass murder, rape, aggravated assault, robbery, burglary, theft, and arson. Both types of sergeant effect are shrunken using the Empirical Bayes procedure described in Section 4. Pearson correlations are provided in black text, with 95% confidence intervals for the correlations, calculated using the Fisher z-transformation, displayed below each correlation. Correlation in the Top 95% is calculated by dropping sergeants below the 5th percentile of non-index effects. Correlation in the Top 50% is calculated by dropping sergeants below the 50th percentile of non-index effects. A best fit-line is given by the dashed red line.

Figure A.12: Relationship Between Serious and Low-Level Officer Effects



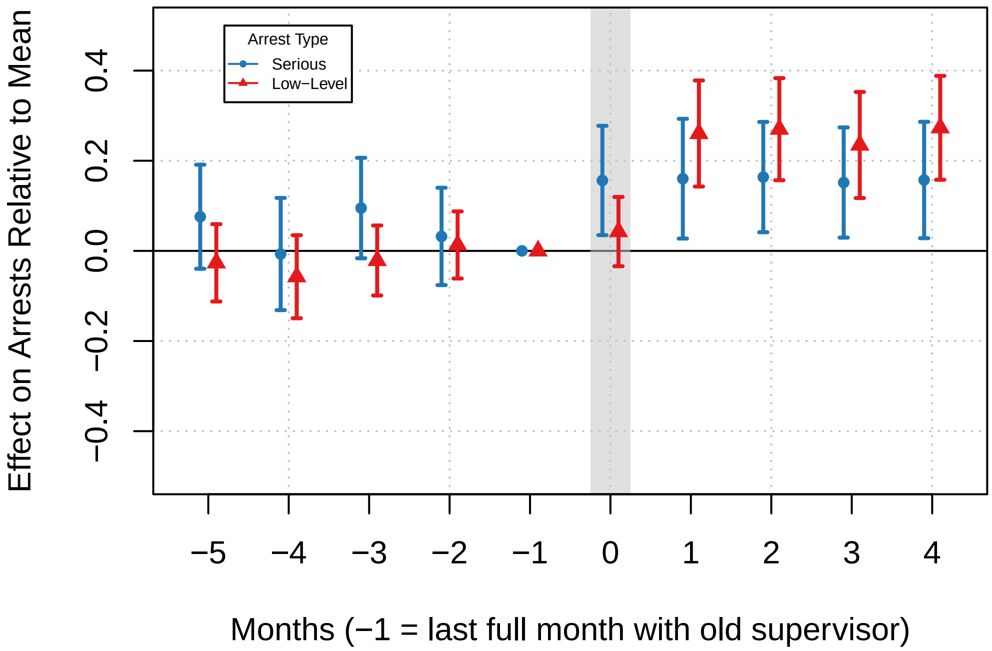
Notes: This figure displays a scatterplot of the relationship between the low-level officer effects and serious officer effects. Each point represents an officer. Low-level (serious) effects describe officer effect on arrests for low-level (serious) crimes, defined as described in Section 3. Both types of officer effect are shrunken using the Empirical Bayes procedure described in Section 4. Pearson correlations are provided in black text, with 95% confidence intervals for the correlations, calculated using the Fisher z-transformation, displayed below each correlation. Correlation in the Top 95% is calculated by dropping officers below the 5th percentile of low-level effects. Correlation in the Top 50% is calculated by dropping officers below the 50th percentile of low-level effects. A best fit-line is given by the dashed red line.

Figure A.13: Impact of Sergeant Effects on Conviction Rates



Notes: This figure plots the change in officer conviction rates that results from increasing the low-level and serious sergeant effects by one standard deviation. I calculate the change in conviction rate in two steps, as described in Section 5. I first regress the number of total and convicted arrests separately on the standardized low-level and serious sergeant effects, along with the model controls as in equation 7. Then, for each of the two sergeant effects, I add the regression coefficients for convicted and total arrests to their respective sample means and take the difference between this ratio and the ratio of the means. For each estimate, I calculate a 95% confidence interval using a bootstrap with 100 resamples. I do this procedure separately for all arrests ('Overall Conviction Rate), serious arrests ('Serious Conviction Rate), and low-level arrests ('Low-Level Conviction Rate').

Figure A.14: Event Study by Arrest Type



Notes: This figure plots event study coefficients for equation 4, separately for models that use serious arrests (defined as index arrests as well as domestic violence, fraud, simple assault, and DUI) and low-level arrests as the dependent variable. Serious arrest results are given in blue and low-level arrest results are given in red. Month -1, the last full month that the officer spends with the old sergeant, is the reference month in all specifications. For each severity level, the effects are divided by the average number of the corresponding category of arrests in month -1. The model is estimated using the event study data that are balanced on [-5, 4].

B Additional Tables

Table B.1: Event-study around sergeant switches

k	Total Arrests (1)	Serious Arrests (2)	Low-Level Arrests (3)
-5	0.0020	0.0732	-0.0712
	(0.1366)	(0.0568)	(0.1175)
-4	-0.1609	-0.0067	-0.1542
	(0.1477)	(0.0613)	(0.1261)
-3	0.0347	0.0919*	-0.0572
	(0.1265)	(0.0549)	(0.1064)
-2	0.0664	0.0310	0.0354
	(0.1191)	(0.0532)	(0.1020)
0	0.2661**	0.1510**	0.1151
	(0.1252)	(0.0597)	(0.1052)
1	0.8550***	0.1548**	0.7002***
	(0.1779)	(0.0654)	(0.1611)
2	0.8843***	0.1583***	0.7260***
	(0.1649)	(0.0602)	(0.1551)
3	0.7782***	0.1466**	0.6316***
	(0.1780)	(0.0602)	(0.1612)
4	0.8862***	0.1520**	0.7342***
	(0.1735)	(0.0636)	(0.1578)
Observations	12,770	12,770	12,770
Y mean	3.6488	0.89742	2.7514
Pre-trends F stat	0.8473	1.5706	0.6234
p-value	0.8473	0.1791	0.6458

Notes: This table presents the event-study coefficients used to make Figures 4 and A.14. The regressions use switching events that are balanced around 5 sample months prior to the move and 4 sample months after the move. Month -1 is the reference point and the switch occurs at in month 0. The pre-trends F statistic is calculated from an F-test of joint significance of the coefficients for which $k < -1$. Standard errors are clustered at the officer level.

Table B.2: Trends in crime do not predict changes in sergeant effects

	All Months		Months With Movers	
	$E[\Delta\hat{\psi}_{out}]$	$E[\Delta\hat{\psi}_{in}]$	$E[\Delta\hat{\psi}_{out}]$	$E[\Delta\hat{\psi}_{in}]$
	(1)	(2)	(3)	(4)
$Log(911Calls)_{-1}$	-0.0162	0.0014	0.0713	0.1357
	(0.0467)	(0.0483)	(0.1654)	(0.1482)
$Log(911Calls)_{-2}$	-0.0043	0.0173	-0.0571	0.0140
	(0.0459)	(0.0570)	(0.1871)	(0.1764)
$Log(911Calls)_{-3}$	0.0063	-0.0594	-0.0722	-0.3354*
	(0.0552)	(0.0479)	(0.1709)	(0.1770)
$Log(911Calls)_{-4}$	-0.0398	0.0148	-0.0915	0.0855
	(0.0515)	(0.0543)	(0.1877)	(0.1545)
$Log(911Calls)_{-5}$	0.0372	-0.0460	0.2426	-0.0737
	(0.0391)	(0.0473)	(0.1723)	(0.1704)
Observations	5,525	5,525	1,387	1,424
Y mean	0.00738	0.00401	0.02938	0.01555
Joint F p-value	0.91938	0.62787	0.79073	0.38557
Sector-Watch fixed effects	✓	✓	✓	✓

Notes: This table examines the correlation of crime trends with sergeant switches. Regressions are performed at the unit by month level. For each unit and month, I calculate the average change in sergeant effects separately for officers who move in and out of the unit in that month. The table reports results for a regression of the sergeant effect changes on the natural logarithm of 911 calls originating in the unit's sector-watch in each of the 5 months prior to the focal month. Regressions include fixed effects for the sector-watch. Dependent variables are the average change in the sergeant effects for out-movers (columns 1 and 3) and in-movers (columns 2 and 4). Out-movers are officers who leave the unit and in-movers are officers who join the unit in a given month. Columns 1 and 2 use all monthly observations for each unit. Columns 3 and 4 only use the monthly observations in which at least one out-move (column 3) or one in-move (column 4) occurs. Standard errors are clustered at the sector-watch level.

Table B.3: Analysis of Variance

	Arrests				
	(1)	(2)	(3)	(4)	(5)
R^2	0.166379	0.516861	0.527611	0.202556	0.623947
Adjusted R^2	0.164421	0.497526	0.505139	0.195091	0.559736
Controls	✓	✓	✓	✓	✓
Officer FE		✓	✓		
Sergeant FE			✓	✓	
Sergeant-by-Officer FE					✓
Observations	49,923	49,923	49,923	49,923	49,923

Notes: This table reports R^2 and adjusted R^2 for models that vary the included fixed effects. Controls include a second degree polynomial of officer tenure, and sector-watch and day-off group fixed effects.

Table B.4: Sergeant Effects by Crime Type

	Serious Crimes			Low-Level Crimes		
	Domestic Violence	Theft	DWI	Drugs	Warrants	Disorderly Conduct
	(1)	(2)	(3)	(4)	(5)	(6)
Low-level Sergeant Effect	-0.0161**	-0.0025	-0.0013	0.1713***	0.1772***	0.1076***
	(0.0080)	(0.0053)	(0.0061)	(0.0238)	(0.0203)	(0.0134)
Serious Sergeant Effect	0.1687***	0.0382***	0.0352***	-0.0360***	0.0755***	0.0350***
	(0.0082)	(0.0064)	(0.0063)	(0.0139)	(0.0148)	(0.0093)
Baseline Controls	✓	✓	✓	✓	✓	✓
Observations	49,923	49,923	49,923	49,923	49,923	49,923
Y mean	0.45308	0.13927	0.10885	0.31138	0.76578	0.40903

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7, with arrests for the three most frequent serious and low-level crimes as the outcome variables. Serious and low-level crimes are mutually exclusive categories. However, within serious and low-level crimes, an arrest may fall under multiple different criminal charges. The low-level (serious) sergeant effect is given by the standardized Bayes-shrunken sergeant effect on low-level (serious) arrests. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table B.5: Sergeant Effects on Arrests by Interaction Source

	Officer Initiated Arrests	Call Initiated Arrests
	(1)	(2)
Low-level Sergeant Effect	0.4451***	0.2561***
	(0.0367)	(0.0237)
Serious Sergeant Effect	0.0339	0.2373***
	(0.0244)	(0.0196)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	1.6951	2.0930

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7 using officer-initiated and call-initiated arrests as outcomes. An arrest is call-initiated if it can be linked to a 911 call in CAD. Otherwise, it is considered officer-initiated. The low-level (serious) sergeant effect is given by the standardized Bayes-shrunken sergeant effect on low-level (serious) arrests. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table B.6: Sergeant Effects on Call Activity

	Calls Answered	Arrest Probability
	(1)	(2)
Low-level Sergeant Effect	1.243***	0.0042***
	(0.4156)	(0.0005)
Serious Sergeant Effect	2.301***	0.0004
	(0.3562)	(0.0004)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	73.542	0.06244

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7 using various measures of call activity as outcome variables. In column 1, calls answered refer to calls in which the officer is recorded as being at the scene in CAD. In column 2, arrest probability at calls is the proportion of calls that an officer answers that result in arrest. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table B.7: Sergeant Effects and Other Activity

	Use of Force Incidents	Complaints
	(1)	(2)
Low-level Sergeant Effect	0.0203***	0.0040
	(0.0043)	(0.0029)
Serious Sergeant Effect	0.0059*	-0.0019
	(0.0034)	(0.0025)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	0.13813	0.02398

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7 using use of force incidents and complaints as outcome variables. The low-level (serious) sergeant effect is given by the standardized Bayes-shrunken sergeant effect on low-level (serious) arrests. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table B.8: Drug Arrest Types

	Possession Arrests	Distribution Arrests
	(1)	(2)
Low-level Sergeant Effect	0.1543***	0.0163***
	(0.0210)	(0.0041)
Serious Sergeant Effect	-0.0314**	-0.0051***
	(0.0132)	(0.0018)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	0.28955	0.01839

Notes: The table reports results for a regression of possession (column 1) and distribution (column 2) arrests on low-level and serious sergeant effects, along with baseline model controls from equation 7, which include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. I classify a drug arrest as "possession" if the charge description only mentions possession and not sale or manufacturing. If the charge description mentions sale or manufacturing, then the drug arrest is classified as "distribution," so that the two categories are mutually exclusive. The effect sizes for all drug arrests are 0.1713 for low-level sergeant effects and -0.0361 for serious sergeant effects, taken from Table B.4. Standard errors are clustered at the officer level. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table B.9: Sergeant impacts by race

	Black Arrests	Hispanic Arrests	White Arrests
	(1)	(2)	(3)
Low-level Sergeant Effect	0.4172***	0.1469***	0.1370***
	(0.0333)	(0.0157)	(0.0127)
Serious Sergeant Effect	0.1150***	0.0966***	0.0519***
	(0.0231)	(0.0154)	(0.0100)
Observations	49,923	49,923	49,923
Y mean	1.9291	1.0082	0.80869

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7, using arrests by race as the outcome variables. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure.. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table B.10: Sergeant Effects and Conviction Rates

	Difference in Conviction Ratio	Difference in Serious Conviction Ratio	Difference in Low-Level Conviction Ratio
	(1)	(2)	(3)
Low-level Sergeant Effect	0.0056**	0.0097*	0.0189***
	(0.0022)	(0.0051)	(0.0138)
Serious Sergeant Effect	0.0089***	-0.0040	-0.0079***
	(0.0019)	(0.0039)	(0.0230)
Baseline Controls	✓	✓	✓
Observations	49,923	49,923	49,923
Mean ratio	0.2046	0.4282	0.1322

Notes: This table reports the changes in officer conviction rates that results from increasing the low-level and serious sergeant effects by one standard deviation, which are plotted in Figure A.13. I calculate the change in conviction rate in two steps. I first regress the number of total and convicted arrests separately on the standardized low-level and serious sergeant effects, along with the baseline controls, which include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Then, for each of the two sergeant effects, I add the regression coefficients for convicted and total arrests to their respective sample means and take the difference between this ratio and the ratio of the means. For each estimate, I calculate a 95% confidence interval using a bootstrap with 100 resamples. I do this procedure separately for all arrests ('Overall Conviction Rate), serious arrests ('Serious Conviction Rate), and low-level arrests ('Low-Level Conviction Rate'). $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table B.11: Sergeant effects and officer overtime activities

	Overtime Calls	Overtime Low-level Arrests	Overtime Serious Arrests
	(1)	(2)	(3)
Low-level Sergeant Effect	0.4808***	0.0312***	-0.0011
	(0.1386)	(0.0057)	(0.0031)
Serious Sergeant Effect	0.4982***	0.0103*	0.0219***
	(0.1238)	(0.0056)	(0.0037)
Baseline Controls	✓	✓	✓
Observations	49,923	49,923	49,923
Y mean	3.7242	0.11327	0.04926

Notes: This table presents results for a regression of overtime calls and arrests on low-level and serious sergeant effects, along with standard model controls from equation 7, which include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure.. I call a call or arrest overtime if it occurs outside of the officer's shift hours listed in the assignments data. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

C Constructing Sergeant Assignments

The Computer Aided Dispatch (CAD) system used by the Dallas Police Department stores assignment indicators for every sworn employee who is assigned to a call. These assignment indicators are known internally as "element numbers." Element numbers are assigned every day to each separate patrol car and describe the watch and beat assignment of the car. Beats are smaller geographic sectors within patrol sectors that individual officers are assigned to patrol.

Watches are described by letters A-F, where A/B/C denote overnight/day/evening watches and D/E/F are variants for day/overnight/evening that allow for multiple units to be assigned to one beat at a time depending on department needs. Beats are given by a 3-digit numeric. Thus, an example of an element number within CAD is A135, which means that the officer is working the overnight shift patrolling beat 135. The first and second digits of the beat code identify the sector in which the beat is located. Returning to the previous example, beat 135 is part of sector 130.

Sergeants are given element numbers that denote the sector and watch to which they are assigned. The sergeant for an officer with the element number A135 has the element number A130. In the case of variant units within a sector, there will be one sergeant in charge of each unit. That is, an officer in the variant overnight unit E135 would have a sergeant with the element number E130. I use this pattern in the element numbers to identify the most common sector-watch assignments for officers and sergeants within each month of the data, as described in Section 3.

The assignments that I construct exclude officer spells in specialty patrol units whose element number does not match a geographic sector. Based on conversations with DPD, these units perform distinct duties from regular patrol officers, as evidenced by the fact that they are not assigned to specific geographic beats. DPD did not provide me with the specific details of the job duties related to these assignments for reasons related to officer safety. But, consistent with a separate and distinct role, officers in these specialty units exhibit significant more variation in arrests than regular patrol officers and sergeants cannot be reliably identified for these units.

D Empirical Bayes Shrinkage

The raw sergeant fixed effects are estimated with error. Suppose that the estimates are given by:

$$\hat{\psi}_j = \psi_j + \epsilon_j, \tag{1}$$

where $\psi_j \sim \mathcal{N}(0, \sigma_\psi^2)$, $\epsilon_j \sim \mathcal{N}(0, \sigma_{\epsilon_j}^2)$, and ψ_j and ϵ_j are independently distributed across the population of 347 sergeants. The mean of the sergeant fixed effects is 0 by construction, since the true mean is unidentified in the model. Under these distributional assumptions, we have that

$$\hat{\psi}_j | \psi_j \sim \mathcal{N}(\psi_j, \sigma_\epsilon^2). \tag{2}$$

Hence, each of the fixed effects are unbiased estimates of supervisor j's effect, as is the case under the identifying assumptions laid out in Section 4. As shown by Morris (1983), one can construct a more efficient estimator of ψ_j using the posterior mean of ψ_j conditional on the estimate $\hat{\psi}_j$:

$$E[\psi_j | \hat{\psi}_j] = \lambda_j \hat{\psi}_j, \tag{3}$$

where $\lambda_j = \frac{\sigma_\psi^2}{\sigma_\psi^2 + \sigma_{\epsilon_j}^2}$. I estimate the shrinkage factor $\hat{\lambda}_j$ by bootstrapping the estimation of equation 1. For each supervisor j, I obtain bootstrap estimates of the fixed effect $\hat{\psi}_j^k$, where k= 1,...,1000. I estimate the error variance of each $\hat{\psi}_j$ using the sample variance of the bootstrap distribution: $\hat{\sigma}_{\epsilon_j}^2 = \frac{1}{k-1} \sum_{k=1}^{1000} (\hat{\psi}_j^k - \bar{\hat{\psi}}_j^k)^2$. I then estimate $\hat{\sigma}_\psi^2$ using the variance estimator proposed by Morris (1983):

$$\hat{\sigma}_\psi^2 = \frac{\sum W_j (\hat{\psi}_j^2 - \hat{\sigma}_j^2)}{\sum W_j}. \tag{4}$$

For my main estimates, I use weights $W_j = 1$, so that the estimate takes the form:

$$\hat{\sigma}_{\psi}^{2} = Var(\hat{\psi}_j) - E_j(\hat{\sigma}_j^2). \quad (5)$$

One can also use the weights proposed by Morris (1983): $\frac{1}{\hat{\psi}_j^2+\hat{\sigma}_j^2}$, which requires estimating $\hat{\sigma}_{\psi}^{2}$ iteratively by first plugging in a guess of the across-supervisor variance and calculating as in equation 5 until the values are sufficiently close. In unreported results, I find that using this weighted estimate yields a similar shrinkage factor.

It is also possible to estimate the variance components directly from the regression residuals using a method proposed by Guarino et al. (2015) and implemented in the policing context by Weisburst (2024). In unreported results, I find that this method produces similar, albeit slightly less conservative shrinkage factors.

E Predicting Sergeant Effects

To what extent are sergeant effects mediated by observable characteristics that are determined *before* someone has been promoted to sergeant? I examine variations in sergeant effect distributions across four observable pre-promotion characteristics: race, gender, age at the time of their promotion exam, and the score achieved on the promotion exam.[1] Since I only observe exam scores beginning with the 2012 round of tests, I limit my sample to the 202 sergeants who were promoted from these exams. I call sergeants "older" if they were above the average across all exams and I call them "above-average scorers" if they were above the average score for their particular exam.

In Figure E.1a, I present densities separately by exam score. I observe a notable difference between high and low scorers in the distributions of both serious and low-level effects. For low-level effects, the distribution of above-average scorers is shifted to the left of that for below-average scorers, with a Kolmogorov-Smirnov test indicating that this difference is statistically significant (p-value = 0.034). Thus, individuals who score below average on the promotional exams tend to induce more low-level arrests than those who score above average. These differences are meaningful considering exam score is the primary determinant of promotion. The exams test for knowledge of department procedures and aptitude within relevant supervisory situations. Thus marginal promotees — who are barely promoted by virtue of their low exam score — are more likely to value low-level arrests.

In Figure E.1b, I also find evidence of a statistically significant difference in the distribution of serious effects between high and low exam scorers (Kolmogorv-Smirnov p-value = 0.013). In contrast to low-level effects, the differences in serious effects are not driven by a monotonic shift in one direction. Instead, high scorers exhibit a wider distribution of serious effects than low scorers. The distribution for low scorers is more concentrated

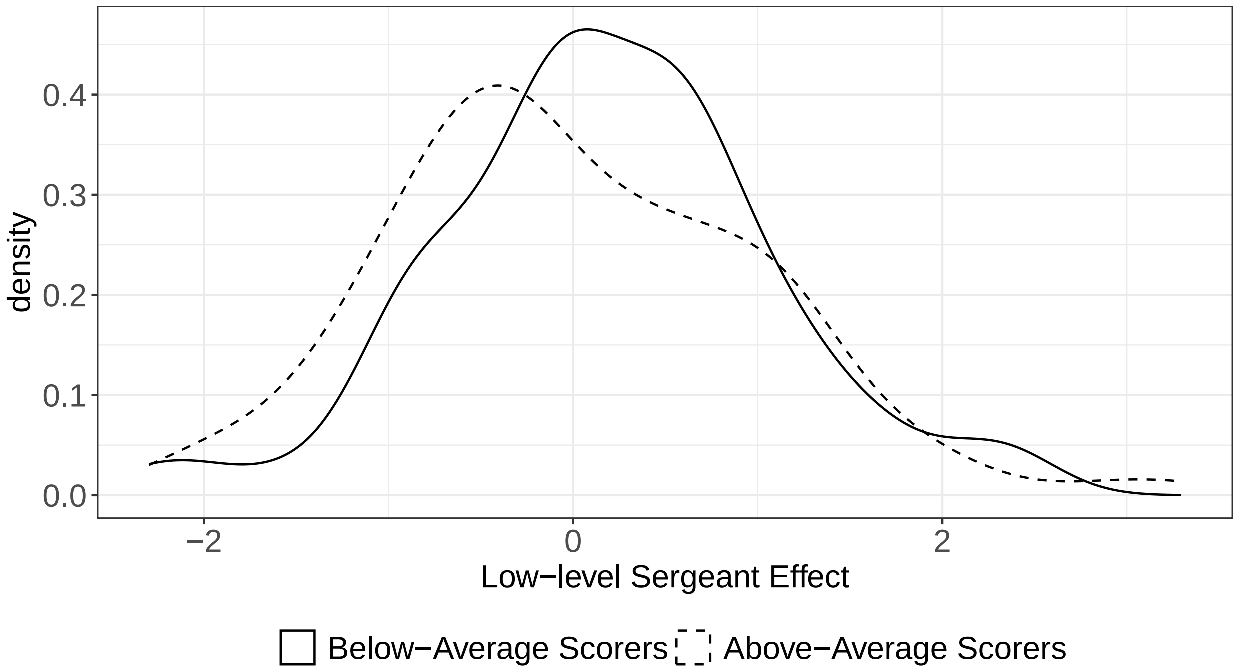
[1] The promotion exam is given once every few years to officers with at least 5 years of experience who want to take part in the promotion process. Exam-takers are ranked according to their performance and, with a few rare exceptions, promoted in order of their exam ranking when openings arise. The exam has both a written and oral component. For this study, I obtained data on the aggregate exam scores for all sergeants promoted during the sample period, accounting for 58% of sergeants in my sample.

around 0. These findings suggest that high scorers are significantly more heterogeneous than low-scorers, at least in terms of their effects on serious enforcement.

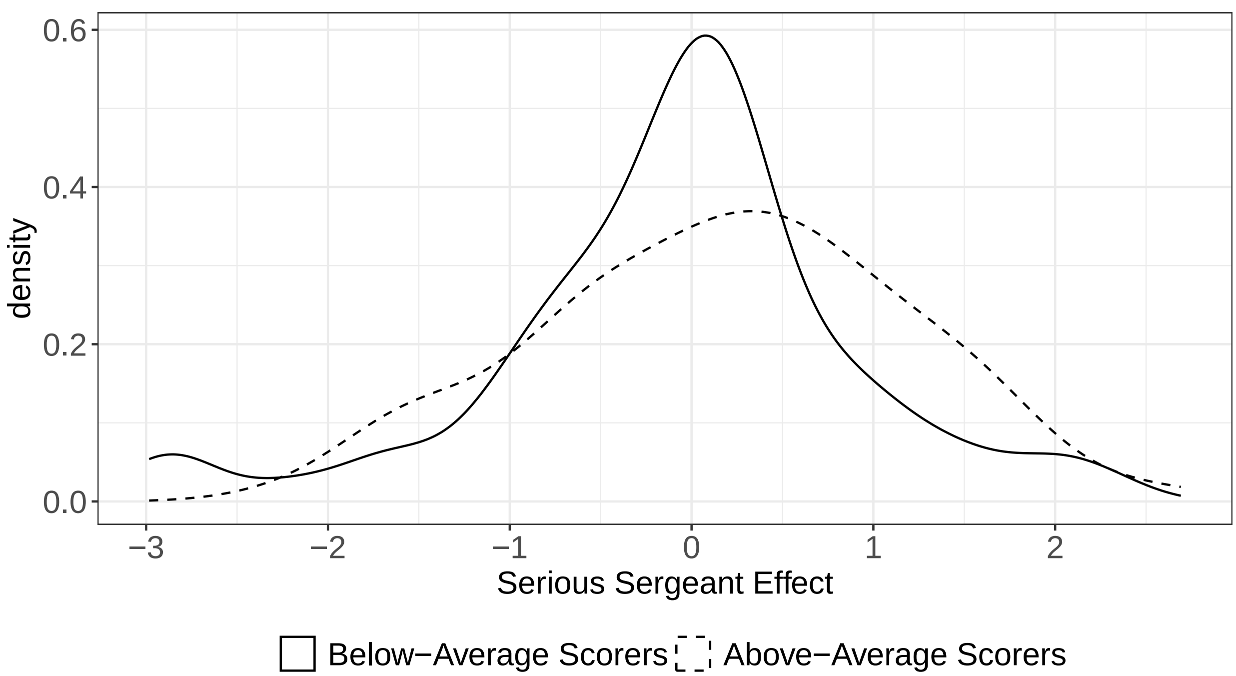
I also analyze the empirical densities of low-level and serious effects separately by race (Figures E.3b, E.3a), gender (Figures E.2b, E.2a), and age groups (Figures E.4b, E.4a). I do not find evidence of significant differences in the distribution of either sergeant effect along each of these three observable dimensions. Kolmogorov-Smirnov tests for the equality of the distributions generate p-values that are well-above standard significance thresholds. However, I find evidence that low-level officer effects differ by race (Figure E.5), consistent with the previous literature (e.g. Ba et al., 2021). Thus, it appears that variation in officer enforcement effects do not translate into similar variation in sergeant effects.

Figure E.1: Sergeant Effects Distributions by Promotional Exam Performance

(a) Low-level effect distribution



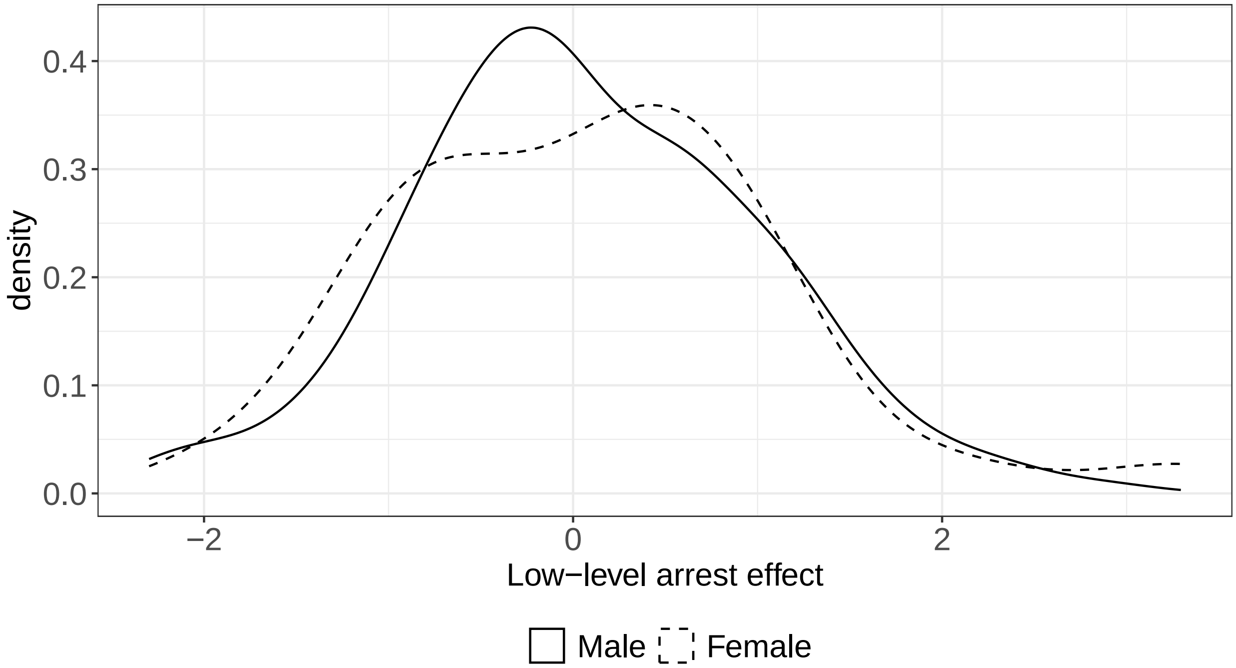
(b) Serious effect distribution



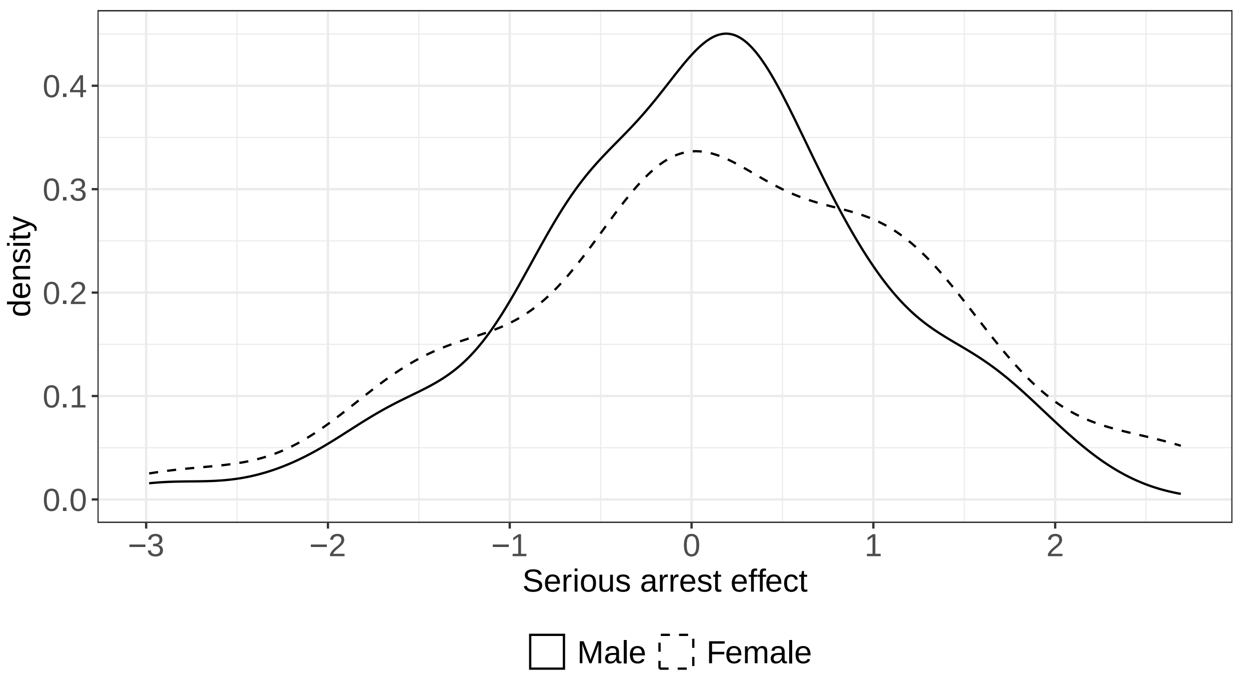
Notes: These figures plot distributions of standardized sergeant effects separately by the sergeant's performance on the promotional exam they took to attain the rank of sergeant. I define above-average scorers as sergeants who score above the average for their exam and below-average scorers as those who score below their exam-specific average. Kolmogorov-Smirnov p-value for low-level effects = 0.0343. Kolmogorov-Smirnov p-value for serious effects = 0.0128.

Figure E.2: Sergeant Effects Distributions by Gender

(a) Low-level effect distribution



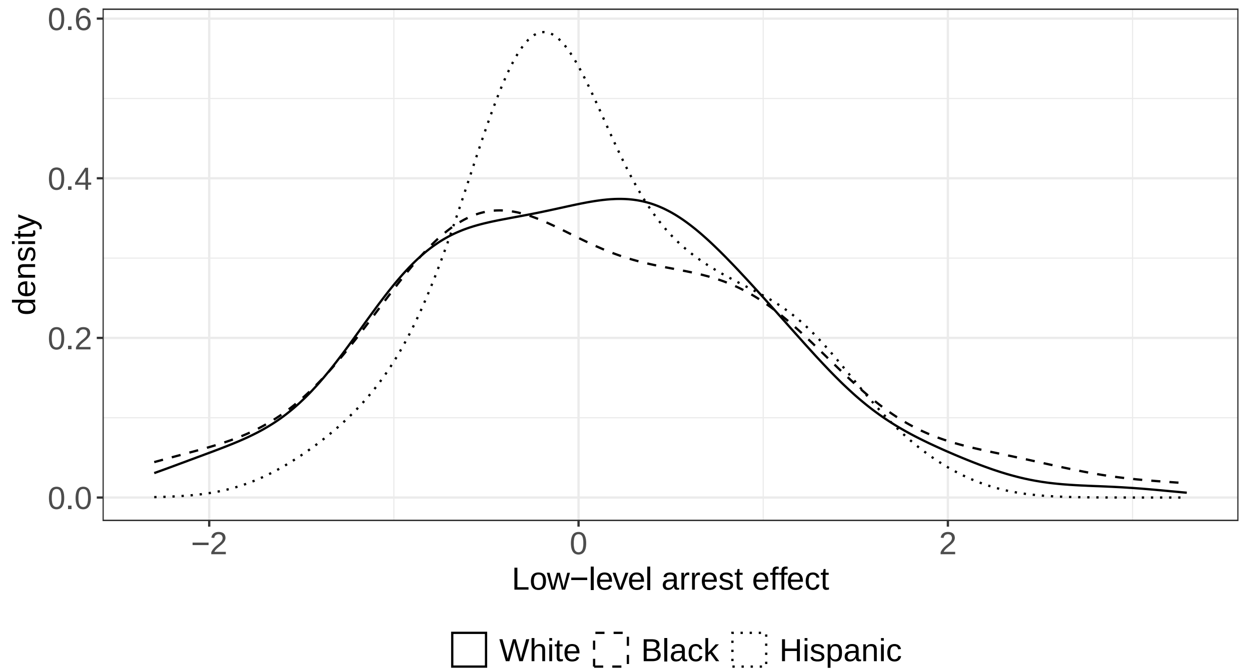
(b) Serious effect distribution



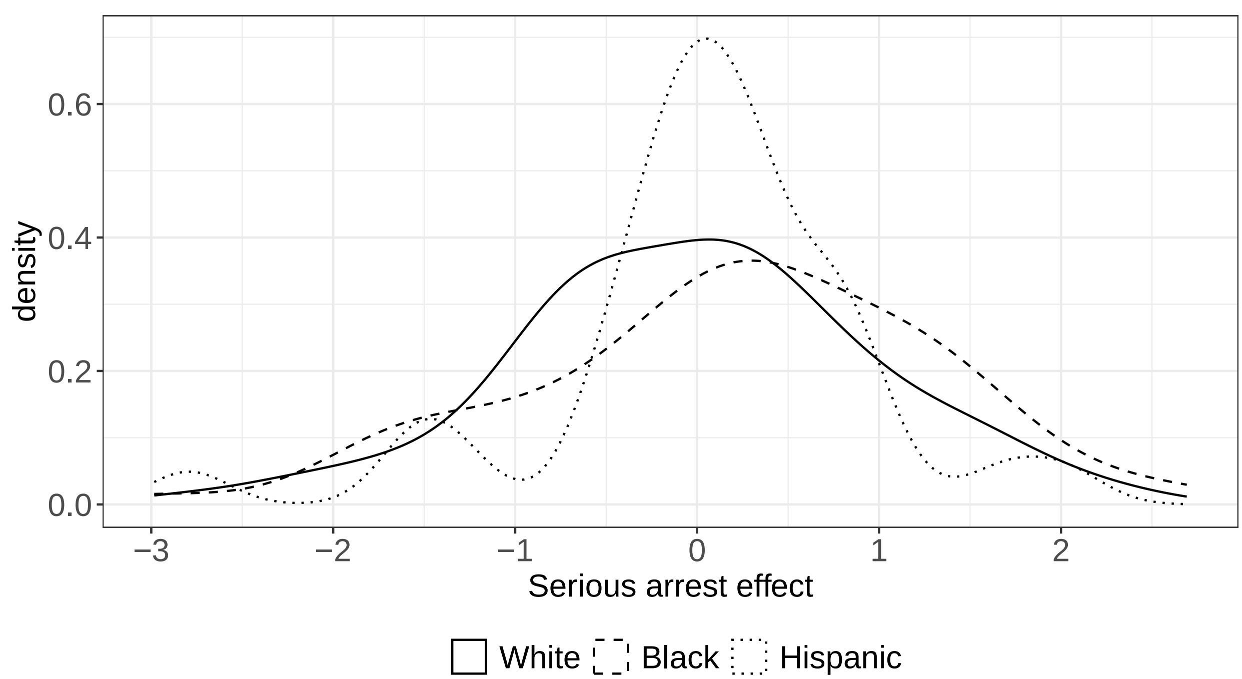
Notes: These figures plot distributions of standardized sergeant effects separately by sergeant gender. Kolmogorov-Smirnov p-value for low-level effects = 0.817. Kolmogorov-Smirnov p-value for serious effects = 0.415.

Figure E.3: Sergeant Effects Distributions by Race

(a) Low-level effect distribution



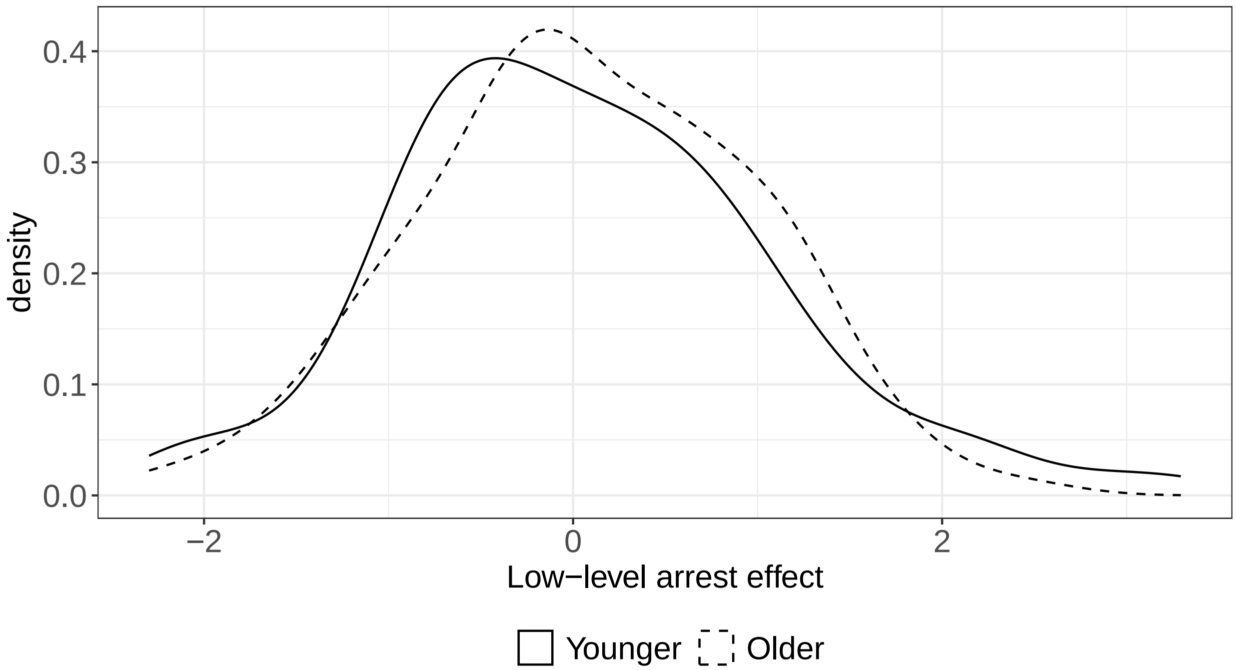
(b) Serious effect distribution



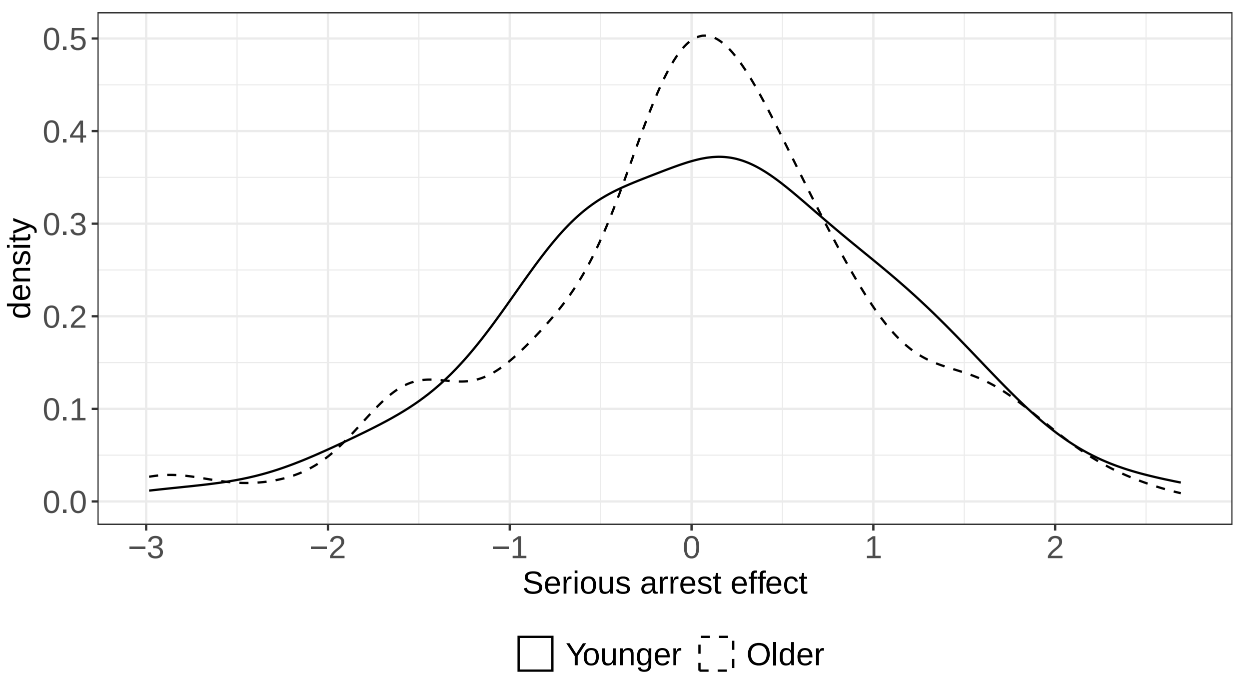
Notes: These figures plot distributions of standardized sergeant effects separately by sergeant race. For low-level effects: Black/Hispanic Kolmogorov-Smirnov p-value = 0.218. Black/White Kolmogorov-Smirnov p-value = 0.241. White/Hispanic Kolmogorov-Smirnov p-value = 0.855. For serious effects: Black/Hispanic Kolmogorov-Smirnov p-value = 0.239. Black/White Kolmogorov-Smirnov p-value = 0.219. White/Hispanic Kolmogorov-Smirnov p-value = 0.130.

Figure E.4: Sergeant Effects Distributions by Age During Promotional Exam

(a) Low-level effect distribution

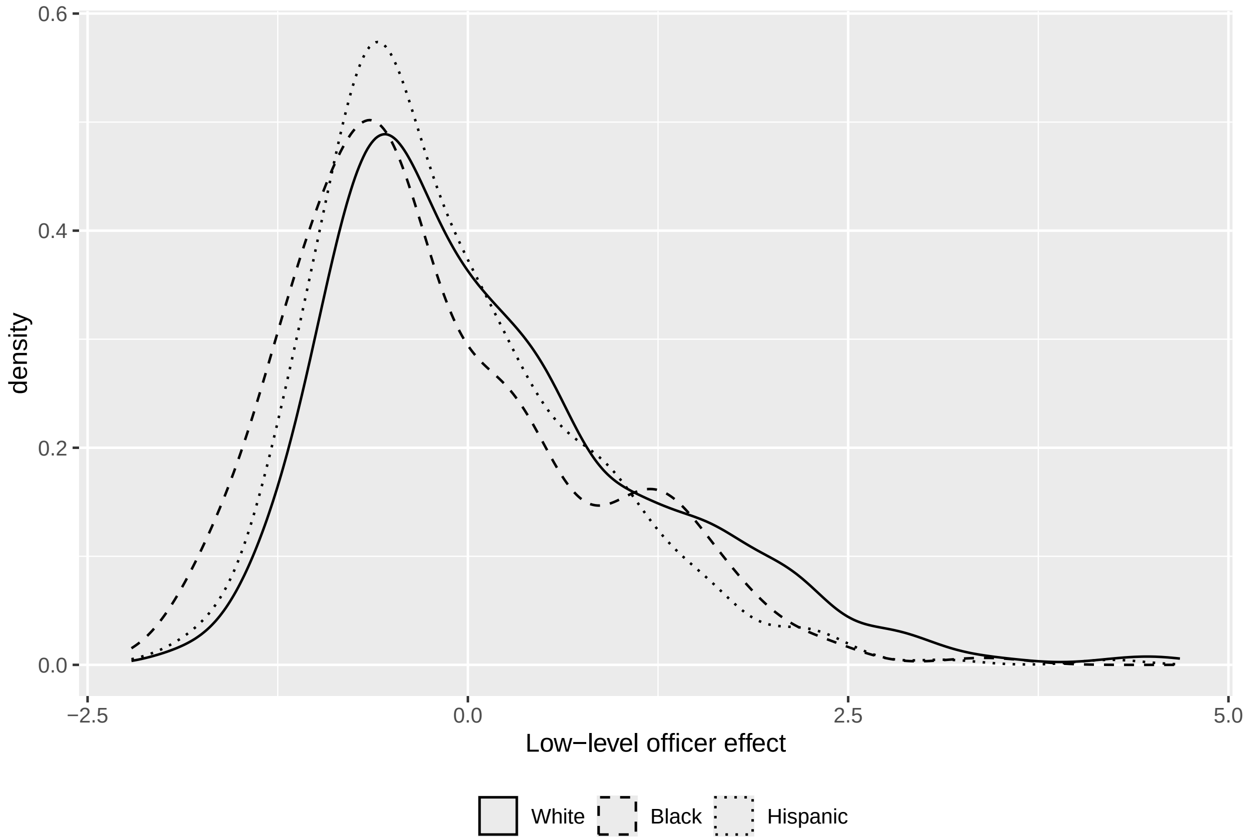


(b) Serious effect distribution



Notes: These figures plot distributions of standardized sergeant effects separately by age at the time of the promotional exam. Older sergeants are those who are above the average age across all exams in my sample, while younger sergeants are below the average. For low-level effects: Kolmogorov-Smirnov p-value = 0.587. For serious effects: Kolmogorov-Smirnov p-value = 0.587.

Figure E.5: Low-level Officer Effects by Race



Notes: This figure plots the distribution of low-level officer effects by officer race. Black/Hispanic Kolmogorov-Smirnov p-value = 0.047. Black/White Kolmogorov-Smirnov p-value = 0.000. White/Hispanic Kolmogorov-Smirnov p-value = 0.003.

References

Ba, B. A., Knox, D., Mummolo, J., and Rivera, R. (2021). The role of officer race and gender in police-civilian interactions in chicago. *Science*, 371:696–702.

Guarino, C. M., Maxfield, M., Reckase, M. D., Thompson, P. N., and Wooldridge, J. M. (2015). An evaluation of empirical bayes's estimation of value-added teacher performance measures. *Journal of Educational and Behavioral Statistics*, 40:190–222.

Morris, C. N. (1983). Parametric empirical bayes inference: Theory and applications. *Journal of the American Statistical Association*, 78:47.

Weisburst, E. K. (2024). Whose help is on the way? *Journal of Human Resources*, 59:1122–1149.