Investigation into Racial and Sexual Bias against Officers within Chicago Police Departments*

Thu Dong Siddharth Arya Hyuk Jang

February 15, 2024

Racial and Gender Disparities within police departments is an insufficiently investigated phenomenon. This paper replicates an investigation into racial and gender bias among officers within the Chicago Police Department and conducts further investigation into the discrimination faced by minority police officers in regard to the quantity of award nominations they recieve. Although the replication portion finds convincing evidence of such discrimination, our investigations were inconclusive in the matter. In any case, this appears to motivate the police department to counterbalance promotion and advancement to control such disparities.

Table of contents

*Code and data are available at: https://github.com/sid19arya/Police-Award-Nomination-Paper-Reproduction. Replication on Social Science Reproduction platform available at: https://doi.org/10.48152/ssrp-5nym-9g31

1 Introduction

There have been many studies that focus on police-civilian interactions with racial disparities but limited research looking into the racial disparities within the police departments. This paper uses the key metric of internal recognition among Chicago Police Department (CPD) officers and its departmental award nominations. The departmental award nomination is chosen as the measure of performance in this paper as actual promotions are very rare in police departments and would be a limiting choice; only 3 percent of police officers in CPD were promoted to sergeant during the period 2007 to 2015. Using this as the metric of interest, we examine evidence regarding racial and gender disparities within the CPD. We do so through he use of a database of information on CPD officers including the following information: demographics, district assignments, civilian complaints, arrests, award performance, and Tactical Response Reports filings. This analysis will have implications on crime and community safety as a diverse police force may improve policing quality (Miller and Segal 2018).

The dataset was obtained from the American Economic Association (AEA 2022). The data was downloaded, simulated, cleaned, and analyzed using R (R Core Team 2023), Tidyverse (Wickham et al. 2019), Tidyr(Wickham 2021), Janitor(Firke 2020), Dplyr(Wickham et al. 2021), Readr(Wickham and Hester 2021), and Stringr(Wickham 2019). Figures and tables were created using Rstanarm(Gabry et al. 2021), Plm(Gabry et al. 2021), Lmtest(Hothorn et al. 2021), Sandwich(Zeileis 2021), and Stargazer(Marek 2021).

The goal of this paper investigate the evidence of racial and gender-based discrimination faced by minority group police officers - as a reproduction from an extension of Rim, Ba, and Rivera's findings on the same subject (Rim, Ba, and Rivera 2020). The paper starts with a discussion of the data source and explains the methodologies used to replicate the tables and graphs of the original paper. Then, we will reproduce the selected results to confirm their findings and improve the accessibility of the data. The paper will conclude with a discussion of the findings in which we apply the methodologies used to find out the correlation between race, gender, and award performance. This research will help us find out if there are racial or gender disparities among CPD officers internally. The results of the research will discuss if there needs to be counterbalancing in promotion and advancement, considering the racial and sexual disparities.

2 Data

2.1 Data sources

The data of the characteristics of CPD is obtained from the original paper (Rim, Ba, and Rivera 2020). The data is already reasonably clean, we minimally process it further to create the desired Tables and Figures. The dataset contains demographics, rank, tenure, district assignment, awards, arrests, use of force measured by Tactical Response Reports filings, and complaints. The data set is focused on the period 2007 to 2015. For simplicity, we use the dataset the authors of the original paper have made and further clean, and analyze it for replication and further research.

2.2 Summary statistics.

Table 1: Summary Statistics for CPD Officer Data

X	Everyone	White	Black	Male	Female
birth year	1981.51	1982.4	1979.54	1981.74	1980.6
start month	7.36 2011	7.36 2011	7.29 2011	7.37 2011	7.34 2011
complaints	0.45	0.44	0.54	0.47	0.37
arrests	23.45	24.8	21.16	24.3	20.07
award performance	7.05	7.47	5.89	7.27	6.21
trr new	0.66	0.72	0.56	0.74	0.32
observation	1715	840	282	1369	346

Table 1 is the replication of the summary statistics of the 1,715 probationary police officers while they were in the police academy. The table summarizes the baseline characteristics of new police officers from CPD. Characteristics from the original table and the reproduced Table 1 that are the same are birth year, start month, complaints, arrests, and observations. One difference is that the replicated table 1 doesn't have the use of force and training. Instead, it includes award performance and Tactical Response Reports filings. This is because the original paper has used their function which makes the characteristic use of force by using award performance and Tactical Response Reports with different coefficients. Out of the whole data set of CPD the reproduction data specifically uses the data of unit 44 to sample the whole CPD data set.

To produce table 1, we must first extract the subset of CPD in which officers are in unit 44. Once the unit 44 data is extracted, we then merge the characteristic data based on their unique NUID (each unique police officer in CPD). After this is done, we can get the characteristics of "Everyone" in which we sum the characteristics and divide it by the observation which is the number of unique officers (NUID) we have. We round the averaged data to two decimal points

like the original paper has done. The race and gender-specific data is obtained by similar methods, but the only difference is that we extract all the rows in the dataset unique NUID where white == 1, black == 1, male == 1, and female == 1. This then gives us 4 subsets of unique NUID with the characteristics of the four different categories we need to make the replicated summarized statistics. Finally, after we have the 5 data sets ("everyone, white, black, male, female") we merge the data sets into one data frame and rename the columns and rows so that it looks almost identical to the summarized statistics. In between these steps, the start month must be modified for us to find the average start month. This is because the dataset contains the start month as a string like 'Jan', and 'Feb' and we need it to be in digits. Thus we must map the strings based on the month for example Jan == 1, ..., Dec == 12. Once this step is done there will not be any NA that needs to be removed from the cleaned data.

The summary statistics show that the average birth year of unit 44 in CPD is 1981 with an average start year of 2011. Average complaints and arrests were 0.45, and 23.45 respectively with an average award performance of 7.05. Comparing the average of the whole unit 44 to white, black, male, and female, we see some trends. The average civilian complaints black officers were getting was higher than the average by 0.09, while female officers were getting 0.08 less than the average. The average number of arrests among black and female officers was lower than the average by 2.29 and 3.38 respectively. Most importantly, the average award performance of black officers was lower than the total average by 1.16 and for female officers it was 0.84 lower. The total number of officers observed is 1715, out of this there was 840 white officers, 282 black officers, 1369 male officers, and 346 female officers.

3 Results

Table 2: Regression Analysis Summary

	Dependent variable:			
	awd_perf			
	(1)	(2)	(3)	(4)
black	-6.217		-3.801	-4.856
	(0.471)		(0.611)	(0.703)
female:black				3.206
				(1.060)
female		-4.368	-2.366	-3.227
		(0.383)	(0.488)	(0.564)

Control for				
Cohort	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Demographics			Yes	Yes
Policing activities			Yes	Yes
Reference group mean	10.8	10.2	11.73	11.73
Observations	5,775	5,775	4,057	4,057

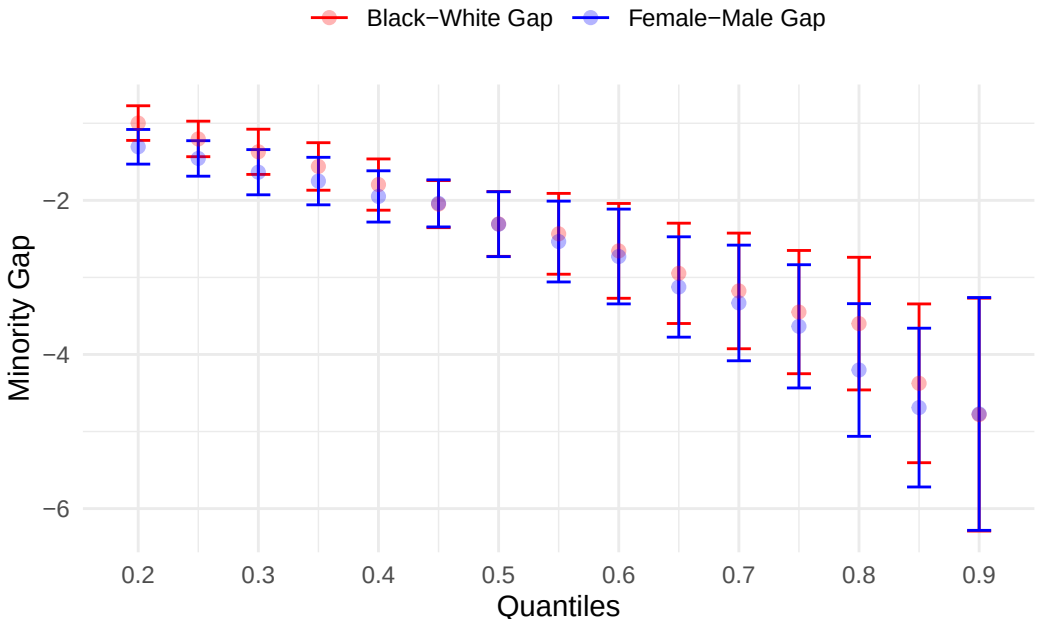

Figure 1: Racial and Gender Discrimination at different Quantile of Award Nominations

3.1 Regression of race and gender on total awards nominations

Table 2 depicts the results of a regression analysis between the demographics of police officers and the amount of awards and nominations received. This table is a replication of an equivalent table in the original paper and shows the regressions of the variables black and female on the total awards nomination with 4 specifications. The result of the replication is slightly different than the result in the paper as the author took an estimate of the sample in the data set without specifying which estimates. Using this method gave the author only 4057 observations out of 5775 observations. However, the results when we use the entire data set are close to the original paper and we come to the same conclusion.

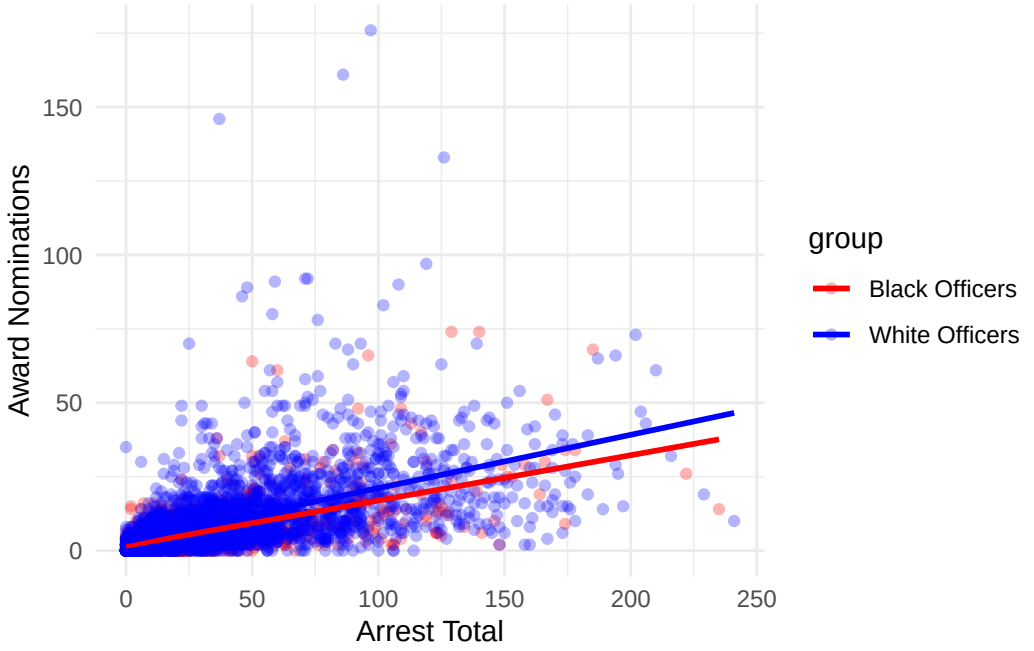

Figure 2: Comparison of Arrest Total and AWD Performance (Black and White)

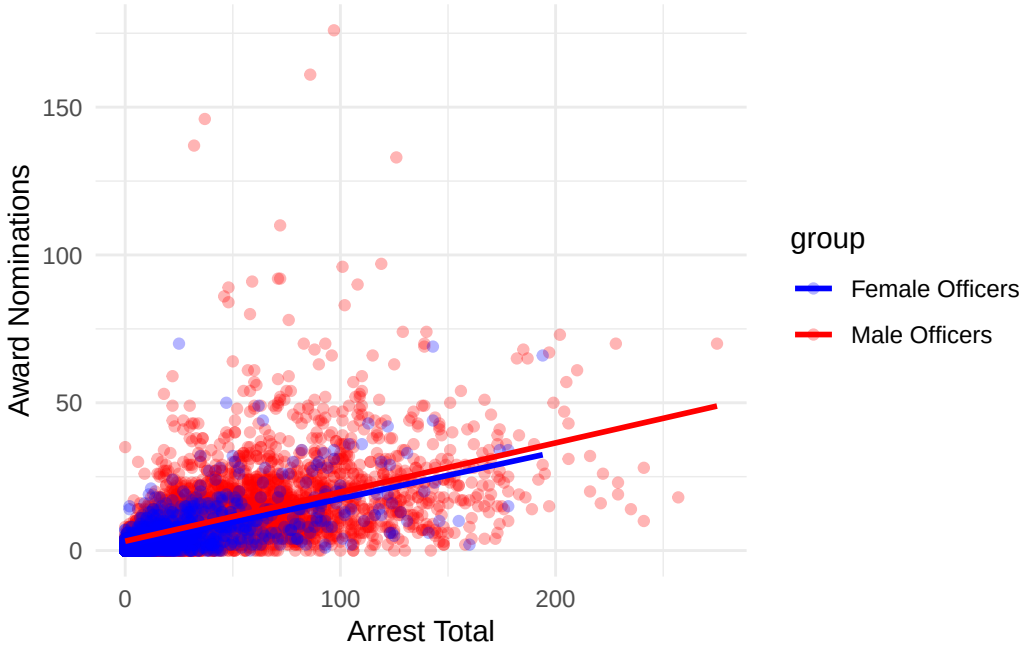

Figure 3: Comparison of Arrest Total and AWD Performance (Male-Female)

Specification (1) shows the regression of race, including Black, Asian, Hispanic, and Native Americans, on awards nominations controlling for cohorts and years, based on the equation below.

$$\text{Award-perf}_{it} = \beta_0 + \beta_1 \text{Black}_{it} + \beta_2 \text{Hisp}_{it} + \beta_3 \text{Asian}_{it} + \beta_4 \text{Natam}_{it} + e_{it}$$

Based on the equation below, specification (2) examines the relationship between gender and total award nominations controlling for cohort and years.

$$\text{Award-perf}_{it} = \beta_0 + \beta_1 \text{female}_{it}$$

Specification (3) then examines the relationship between gender, race, and total award nominations with control of dome confounding variables police activity and demographics. Finally, specification (4) shows the regression of gender and race on awards nominations with the same control variables as the previous specification. This specification specifically includes the interaction terms. The equation will be shown below. The results of this regression will be examined in the discussion section.

$$\text{Award-perf}_{it} = \beta_0 + \beta_1 \text{Black}_i + \beta_2 \text{Female}_i + \beta_3 \text{Black} \times \text{Female}_{it} + \text{other confounding variables}$$

3.2 Bias across different levels of performance

Figure 1, depicts a quantile regression on every quantile between the twentieth and ninetieth percentile of annual award nominations, following the original paper. As the quantile increases, the gap between the number of award nominations between the minority group(black and/or female) and the majority group(while and/or male) gets wider. In other words, in the highest quantiles, the bias against the minority group increases. Specifically, in the twentieth percentile, while black people receive 0.997 fewer awards on average compared to white officers, female officers receive 1.08 fewer awards than their male counterparts. In the fiftyth percentile, the black-white gap is -2.31 and the female-male gap is also approximately -2.31 which means they receive significantly fewer nominations when they perform better. Moreover, in the ninetyth percentile, the recognition for the work of black and female officers reduces even more, as black officers get 4.78 fewer awards than their white counterparts and female officers get -4.77 fewer awards than the male officers.

3.3 Recognition of Work for Officers

Next, building on the research laid out in the original paper, we extend it to investigate a new, connected claim: whether the work of officers of minority groups are less recognized and rewarded for their work. To investigate this claim we observe the correlation between the number of arrests made (a measure of 'work') and the number of award nominations (a measure of reward), for officers of different demographics.

Figure 2 is the scatterplot that illustrates the correlation between the total number of arrests and the total number of awards nominations of 2 groups black and white. Overall, there could be a positive linear relationship between total arrests and total awards nominations. The slope of the line representing the relationship between the number of arrests and the number of awards nominations is steeper for white officers compared to black officers. However, since the differences between the two slopes are marginal and the correlation between total arrests and several awards nominations for black officers and white officers are, respectively, 0.6 and 0.53, which means that there is only a moderate relationship, we were not able to show any evidence of racial bias.

Similarly, we conducted the same analysis for female officers' and male officers's work performances. Figure 3 shows the scatterplots that illustrate the correlation between total arrests and total awards nominations of females and males. Overall, there is a positive linear relationship between the two variables. The slopes of the line representing the relationship between the number of arrests and the number of awards nominations for both groups are approximately parallel which could mean that we can't conclude that there is sexism in the police department.

4 Discussion

4.1 Disparities in police nomination award

Specification (1) displayed the regression of race on the number of awards nominations. Since the data set is panel data, the fixed effect method was used in the analysis to control for some individual-specific characteristics that don't change over time which helped to control for some unobserved factors to address potential endogenity. Specification (1) controls for "cohort" which is defined as all officers who start in the same district in the same quarter and years. Column 1 yields a (b1) of -6.217 which is 17% higher than the original paper results. This result is statistically significant at 1% and indicates that on average, black officers get 6.2 fewer awards than their white counterparts. With the average awards white officers received of 10.8 awards, black officers received 56.7% fewer awards than white officers.

Specification (2) examined the relationship between gender and the number of awards nominations using the fixed effect model, controlling for cohorts and years. Column 2 yields a (b1) of -4.368 which is 21% higher than the original paper results. At a 1% significance level, This

result indicates that on average, female officers get 4.4 fewer awards than male officers. Moreover, since male officers receive 10.2 awards nominations on average, female officers receive around 43% fewer awards than their male counterparts.

Specification (3) shows the effect of race and gender on the number of awards nominations using the fixed effect model, controlling for cohorts and years. To reduce omitted variable bias and address endogeneity, we control for demographics(birth year, tenure) and police activities(total arrests, civilian complaints, and uses of forces). The original paper specified why some of these variables were added as they believe these variables correlated with the number of awards given. For example, older officers are more likely to be nominated because they have more experience in the field. Moreover, authors control for the number of arrests as they believe black officers are more likely to be assigned to black neighborhoods where there are more crimes, in which case they will have more arrests on their records. Therefore, by controlling for these variables, the analysis focuses only on the systemic issues of the internal police department. Specification (3) yields a (b1) of -3.801, which is 6% higher than the original paper's results, and a (b2) of -2.366, which is about 4% lower than the original paper's results. The magnitudes of the coefficient after controlling for additional variables decrease which shows that there were omitted variable biases in the first 2 specifications. The result is statistically significant at 1% and indicates that on average, black male officers get 3.8 fewer awards than their white counterparts. Additionally, with the average awards of officers received of 11.73 awards, black male officers received approximately 32% fewer awards than average white male officers. Specification (3) also stated that there white female officers get 2.4 awards less than white male officers on average which means they received around 20.4% fewer awards than white male officers on average. It also shows that black female officers on average receive 6.2 fewer awards than the white male officers which is 53% fewer awards than their male counterparts.

Specification (4) illustrates the regression of gender and race on the number of awards nominations using the same method and control variables as specification (3). However, an interaction term of black and female variables was added to see how the relationship between black and female affects awards nominations number. Column 4 yields a (b1) of -4.856, a (b2) of -3.227, and a b3 of 3.206 which is, respectively, 5% higher than, 3.4% lower than,5.7% higher than the original paper's results. All coefficients are significant at the 1% level. Specification (4) indicates that on average, black male officers get 4.8 fewer awards than their white male counterparts. Additionally, with the average awards of officers received of 11.73 awards, black male officers received approximately 41% fewer awards than average white male officers. Specification (3) also stated that there white female officers get 3.2 awards less than white male officers on average which means they received around 27% fewer awards than white male officers on average. It also shows that black female officers on average receive 4.8 fewer awards than the white male officers. We witnessed that the number of awards that black officers receive are around the same regardless of their gender.

4.2 Implication of Findings

The issue of systematic racism has been ongoing in the police department in the US. The data set from CPD and the original paper has highlighted the issues of racism and sexism. To extend the findings of the original research, we conducted further analysis to investigate how the officer's police activities, specifically arrest records, affect their awards nominations. The motivation behind our analysis is to see how officers of different races and genders are valued based on mere merit.

Comparing the 2 results, the result from the original paper gives definitive evidence suggesting biases in the police department while our approach shows that there are inconclusive results on the biases if basing awards nominations on merit. These results even though are contradictory, they sparks an interest in further investigation of the police department with different populations and methods to get a more conclusive result.

4.3 Acknowledgment of systematic issues

While it is still uncertain whether or not there are some biases in the internal police department, there are some systematic issues that should be addressed. Firstly, the underrepresentation in leadership of minority groups in the police department is one of the issues that causes biases. The lack of representation of black officers and female officers can lead to bias in valuation, tas assignment, and opportunity for advancement. Moreover, public perception is also an issue that we witness in the law authorities. Negative stereotypes of female officers and Black officers lead to the underrecognition of their work and contributions.

4.4 Weaknesses and next steps

There are multiple weaknesses of the research done throughout this paper. Because Chicago is a segregated city in terms of race and income, black officers are more likely to be assigned to black neighborhoods which have higher crime rates leading to higher award performance. Another is when the average start month is calculated after mapping each month to its numeric value, it is likely that the larger the number, it distorts the average due to the small range [1:12]. This may be the reason why the start month data is all averaged around July. In addition, the research was based on the cleaned data that had been provided from the authors of the original paper. This may limit our conclusions as the original authors have removed some data from the source which may be valuable in inducing our conclusions. For more accurate results it would be recommended to do further research using data sets such as national officer data with similar characteristics.

5 Conclusion

In conclusion, this paper has explored how the replication of the tables and graphs is done, and if there are any disparities among gender or race. This was done using the dataset that was provided in the original paper (Rim, Ba, and Rivera 2020). The original paper has concluded that the analysis presents evidence of supervisors being biased against minority officers. However, further research on the same idea found no evidence of the effort and work of minority officers being less recognized than their counterparts. That being said, the present indication of discrimination should motivate the police department to investigate and find countermeasures for this disparity.

References

Firke, Sam. 2020. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gabry, Jonah et al. 2021. *Rstanarm: Bayesian Applied Regression Modeling via Stan*. https://mc-stan.org/rstanarm/.

Hothorn, Torsten et al. 2021. *Lmtest: Testing Linear Regression Models*. https://cran.r-project.org/package=lmtest.

Marek, Hlavac. 2021. *Stargazer: Well-Formatted Regression and Summary Statistics Tables*. https://cran.r-project.org/package=stargazer.

Miller, Amalia R, and Carmit Segal. 2018. "Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence." *The Review of Economic Studies* 86 (5): 2220–47. https://doi.org/10.1093/restud/rdy051.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rim, Nayoung, Bocar Ba, and Roman Rivera. 2020. "Disparities in Police Award Nominations: Evidence from Chicago." *AEA Papers and Proceedings* 110 (May): 447–51. https://doi.org/10.1257/pandp.20201118.

Wickham, Hadley. 2019. *Stringr: Simple, Consistent Wrappers for Common String Operations*. https://stringr.tidyverse.org/.

Wickham, Hadley et al. 2021. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org/.

Wickham, Hadley. 2021. *Tidyr: Tidy Messy Data*. https://tidyr.tidyverse.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Jim Hester. 2021. *Readr: Read Rectangular Text Data*. https://readr.tidyverse.org/.

Zeileis, Achim. 2021. *Sandwich: Robust Covariance Matrix Estimators*. https://cran.r-project.org/package=sandwich.