

# Tossing Out the “Bad Apples”: Predicting Civilian Complaints Against NYPD Employees

Presented by Alexander Eichler. Co-Authors: Christopher Arnold and Sean Kaukas  
Boston College, Woods College of Advancing Studies, M.S. in Applied Economics



## Introduction

At the moment, America sits divided in more ways than one. That said, due to the murders of George Floyd, Breonna Taylor, Ahmaud Arbery, and others, one of the most important issues at the moment is systemic racism in police forces. While some argue that NYPD employees that murder are just a handful of a “few bad apples,” there are others who contend that those bad apples spoil the whole bunch.

While it may be impossible to predict whether or not a police officer will commit murder against a racial minority, we can try to predict whether or not they will be problematic. On July 26, 2020, ProPublica released New York Civilian Complaint Review Board (CCRB) data that spans the years 1985-2020. The data contains each individual complaint that was filed against an NYPD officer over that time period, and whether or not the CCRB found the complaint to be substantiated. While NYPD unions are fighting against the data being public, ProPublica made the data available for download. Our intent is to use this information to determine what factors influence the likelihood of a police officer (or NYPD employee of higher rank) receiving more valid complaints filed against them.

Precincts across the country should actively be trying to minimize the number of complaints that are filed against their employees because complaints are expensive. The City of Chicago had to spend over \$500 million over the course of ten years due to legal fees and settlements from complaints. So, it would necessarily do New York City and the NYPD well to know which employees are more likely to receive complaints filed against them.

Therefore, our goal is to use the over 33,000 records in the ProPublica dataset to determine what factors influence the likelihood that an NYPD employee will induce substantiated complaints. We will be specifically looking at demographic information of both the employee and the complainant, the number of complaints filed against them, and the types of complaints filed. There are four types of complaints called FADO types, with FADO being an acronym for Force, Abuse of Authority, Discourtesy, and Offensive Language. We suspect that complainant demographic information might have an effect on the number of valid complaints received - for example, an employee who repeatedly receives complaints from black people might be profiling them and could receive more complaints in return. Additionally, we predict that FADO type may have an impact too. If an employee has received numerous complaints for excessive use of force or for abusing their powers, then they might be more likely to continue to do it since they likely have not been fired for it.

## Key Terms

A. **CCRB** - Refers to the Civilian Complaint Review Board of New York City, which reviews every complaint filed against every police officer, decides whether or not the claim is substantiated, and if so, selects a punishment for the officer if necessary.

B. **Substantiated** – A demarcation from the CCRB on a complaint. According to ProPublica: “The alleged conduct occurred and it violated the rules.” Results in punishment for the officer.

C. **Unsubstantiated** – A demarcation from the CCRB on a complaint. According to ProPublica: “The CCRB has fully investigated but could not affirmatively conclude both that the conduct occurred and that it broke the rules.”

D. **Exonerated** – A demarcation from the CCRB on a complaint. According to ProPublica: “The alleged conduct occurred but did not violate the NYPD’s rules, which often give officers significant discretion over use of force.”

E. **FADO Type** - Acronym of the four categories of allegations against officers:

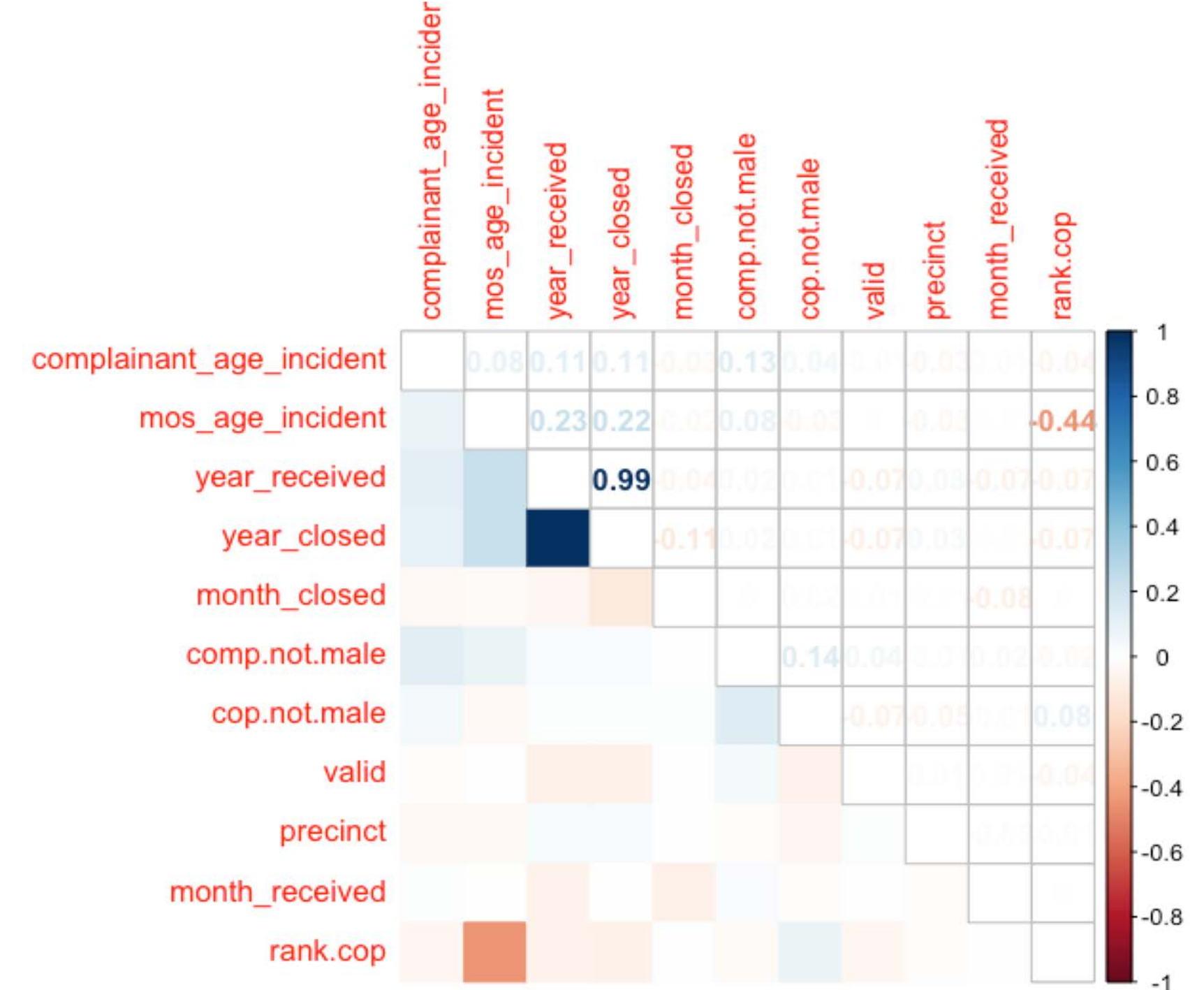
1. **Force** - physical in nature
2. **Abuse of Authority** - manipulation of position of power
3. **Discourtesy** - general verbal attack or gesture
4. **Offensive Language** - verbal attack specific to a marginalized group

## Methodology

The nature of our problem is one that requires the prediction of a discrete number, so model types we will consider will be count models. Overdispersion of the data will suggest that a Negative Binomial model makes the most sense in our situation. A Zero-Inflated model would not make much sense to consider because every employee in the data had a complaint filed against them. While a handful of employees had no valid claims filed against them, it was not a significant enough number to consider a Zero-Inflated model.

Variable	Description
valid	# of valid complaints for a given police officer.
invalid	# of invalid complaints for a given police officer.
invalid.Force	# of invalid complaints of FADO type = Force.
invalid.AbuseAuthority	# of invalid complaints of FADO type = AbuseAuthority.
invalid.Discourtesy	# of invalid complaints of FADO type = Discourtesy.
invalid.OffensiveLang	# of invalid complaints of FADO type = OffensiveLang.
cop.age	Average age of NYPD employee across all their incidents.
cop.white	Binary variable: If NYPD employee ethnicity is "white", then 1, otherwise, 0.
cop.male	Binary variable: If NYPD employee gender is "male", then 1, otherwise, 0.
comp.age	Average age of the complainants across all of the NYPD employee's complaints.
comp.percent.minority	% of NYPD employee's complainants where ethnicity != "white".
comp.percent.not.male	% of NYPD employee's complainants where gender != "male".
rank.cop	Binary variable: If NYPD employee rank = "police officer", then 1, otherwise, 0.

## Variables and Results



**Model 1:** We used backward elimination to get a model with 5 significant variables, as per the table below. The coefficients for the invalid claim counts are consistent with their theoretical effects – a higher number of invalid claims should result in a higher number of valid claims. Comp.percent.not.male was consistent with what we would have expected given previous data exploration, as the majority of complaints were filed by non-males. The coefficient tells us that NYPD employees who have a higher percentage of their complaints coming from non-males have a lower number of valid claims against them. This may indicate some bias from the review board and its rulings.

Variable	Coefficient	Significance
(Intercept)	0.360	***
Invalid.Force	0.032	***
Invalid.Discourtesy	0.049	***
Invalid.AbuseAuthority	0.044	***
Invalid.OffensiveLang	0.066	*
Comp.percent.not.male	-0.345	***

**Model 2:** We took Model 1 and used the total count of invalid claims instead of showing the breakout by FADO type. This model gives us less information, but it is clean and all variables are significant at the .001 level. The coefficients are consistent with Model 1. The coefficient for invalid looks to be roughly the average of the coefficients for invalid.Force, invalid.Discourtesy, invalid.Authority, and invalid.Lang, which makes sense.

Variable	Coefficient	Significance
(Intercept)	0.357	***
Invalid	0.042	***
Comp.percent.not.male	-0.339	***

## Discussion

While both models are simple and intuitive, Model 1 is more useful because it offers up more information. Having the breakout of invalid claims by FADO type allows us to see that offensive language, while it has the highest coefficient, is the least significant. This tells us that if you happened to have two NYPD employees, employee 1 with a large % of their complaints for abuse of authority and employee 2 with a large % of their complaints for offensive language, you could argue that employee 2 is less of a risk for a valid claim than employee 1. Both models compare comparably when measuring against predictions of the training data, while Model 2 is slightly better at predicting complaint number against the test data.

It is clear that the number of valid complaints an NYPD employee accumulates can be reasonably predicted by the number of invalid complaints accumulated and the percent of complainants that are non-male. Three types of complaints (force, discourtesy, and abuse of authority) are significant predictors (at the .001 level), and we can have more confidence in predictions that have a majority of these complaint types. If an NYPD employee's complaints against them came from a large percent of non-male complainants, then we'd predict that the number of valid complaints would be lower. This could be a reflection of review board bias, which we'll expand on later.

As we mentioned in the introduction, the costs associated with processing, reviewing, and settling complaints can be very expensive for police departments. This is a significant motivator for police departments to keep complaints to a minimum, in addition to all social costs. Therefore, having a structure in place to monitor law enforcement employees and anticipate potential valid complaints can help reduce these financial and social costs. The significant predictors that we discussed can be useful inputs to this process.

Regarding review board biases, we cannot assume that the review board is making accurate decisions for each complaint that comes across their desks. In a subsequent regression, we tested the significance of potential biases by building a binary logistic regression model to predict the review board's decision (valid or invalid) of a complaint. To summarize results, it appears that there are a variety of biases that the review board has:

- Complaints filed by minorities are less likely to be ruled valid.
- Complaints filed by non-males are less likely to be ruled valid.
- Complaints filed against non-male NYPD employees are more likely to be ruled valid.
- Complaints filed against police officers (lowest ranking NYPD employee) are more likely to be ruled valid.
- Abuse of authority complaints are most likely to be ruled valid.

## Reference

Derek Willis, E. (2020, July 26). The NYPD Files: Search Thousands of Civilian Complaints Against New York City Police Officers. Retrieved August 16, 2020, from <https://projects.propublica.org/nypd-crb/>