# (Class) Disparities in Policing

Lily Medina

May 08, 2025

**Abstract**

This paper proposes a game-theoretic model to explain class-based disparities in police violence, emphasizing how unequal access to judicial institutions shapes citizen-police interactions and the spatial distribution of officer behavior. Unlike existing accounts of taste-based or statistical discrimination, this model highlights how structural inequalities in complaint mechanisms lead to a self-reinforcing dynamic: poor citizens, facing higher costs and lower probabilities of success when reporting police misconduct, are less likely to denounce abuse. Anticipating minimal risk of punishment, aggressive officers gravitate toward these areas, while non-aggressive officers prefer wealthier neighborhoods where complaints are more credible. The model yields clear comparative statics and multiple equilibria, showing how institutional design and officer assignment policies can either exacerbate or mitigate these disparities. The results offer a theoretical foundation for understanding how structurally embedded biases in access to justice produce class-based disparities in policing—even in the absence of prejudice by police officers.

## 1 Introduction

Giovanni López, a bricklayer from Jalisco, Mexico, was detained on the night of May 4, 2020, outside his home in the Los Olivos neighborhood of Ixtlahuacán de los Membrillos for allegedly not wearing a facemask. The following day Giovannis' relatives learned that he had died in a hospital after being beaten by police officers in a police station. A month later, inspired by the public's response to the 8-minute video that showed Derek Chauvin kneeling on George Floyd in Minneapolis, United States, Giovanni's brother shared footage of Giovanni's arrest. Many Mexicans

responded with outrage and turned to the streets to demonstrate against police brutality. Both government officials and the general public questioned the excessive use of force by the police. The Jalisco State Human Rights Commission ultimately concluded that the police had used excessive force without any evidence or justification for their actions (CEDHJ 2020.)

This case is not an isolated instance of police violence. In the United States, about 1000 civilians are killed by the police every year (Peeples 2020). In a survey of detained people conducted by Mexico's statistical agency (INEGI) in 2016 [1], around 64% of respondents reported having been physically abused during the arrest.

While most incidents of police abuse go unnoticed, those that make the headlines often serve as breaking points to spark discussions about police reform and biases against specific social groups. In Latin America, these debates frequently center on the apparent tendency of officers to use excessive force against poorer individuals. Whether and to what extent class contributes to disparities in policing remains both an academic question and a pressing civic matter.

Drawing from the economic literature on labor market discrimination, scholars studying police violence suggest two primary explanations for disparities in policing across social groups: taste-based and statistical discrimination. The former, taste-based discrimination, posits that individuals may act on preferences to favor or disadvantage certain groups based on traits or characteristics unique to those groups. In the context of class and police violence, this explanation suggests that police violence against the poor arises from greater sympathy for the rich or unfounded prejudices against the poor. The latter, statistical discrimination, hinges on the notion that certain groups are statistically more likely to engage in criminal behavior, leading to heavier policing of these groups and a consequent increase in incidents of police violence.

These two frameworks are grounded in a rich and extensive body of literature on racial discrimination (Gelman, Fagan, and Kiss 2007; O'Flaherty and Sethi 2008; Lerman and Weaver 2014a; McCall 2019a; Knox, Lowe, and Mummolo 2020; Clark et al. 2023; Clark 2024), while class-based explanations have been comparatively underexplored. I propose an alternative mechanism to explain class-based disparities in policing—one that operates independently of assumptions about the

[1] Encuesta Poblacion Privada de la Libertad (ENPOL): https://www.inegi.org.mx/contenidos/programas/enpol/2016/doc/2016_enpol_presentacion_ejecutiva.pdf

behavior of particular groups and is related, though distinct, from police preferences for or against specific class groups.

More precisely, the mechanism this paper proposes is that class-based disparities emerge due to structural class biases embedded in other institutions, which influence policing outcomes independently of the officers' beliefs. The rationale behind this argument is that we would observe disparities in policing tactics across socioeconomic groups because wealthy citizens might be more likely to denounce abuses by the police than poor citizens. Law and society scholars have explained failures in access to justice in terms of resources: individuals of low socioeconomic status might lack the knowledge, financial resources, and time to turn to judicial institutions in the face of an unfair situation (Seron and Munger 1996).

I incorporate this intuition into a game-theoretical model that captures the dynamics between police behavior and citizen engagement. The model considers two types of officers: those who prefer aggressive tactics and those who favor non-aggressive tactics. It also includes parameters representing how easily citizens can access judicial institutions, used as a proxy for class. Through this framework, the model examines the conditions under which residents cooperate with the police and officers opt for non-aggressive tactics.

The findings sugges a self-reinforcing cycle where officers' behavior and residents' engagement influence each other. When access to judicial institutions is poor, residents are less likely to report police misconduct, creating an environment that attracts aggressive officers. On the other hand, when access to judicial institutions is improved, residents are more willing to engage with the police and report misconduct, which discourages aggressive officers and attracts non-aggressive ones.

Building on these insights, an extension of the model endogenizes officers' neighborhood choices, linking disparities in access to judicial institutions with the spatial distribution of officers. The model provides a framework to understand class-based disparities in policing: aggressive officers concentrate in impoverished neighborhoods with weak judicial access, while non-aggressive officers gravitate toward wealthier areas with stronger institutional safeguards.

This paper is structured as follows: it begins with a brief discussion of the existing literature explaining the potential positive relationship between police abuse and low socioeconomic status.

Section 3 introduces and motivates the initial version of the model, which incorporates varying access to judicial institutions. This section includes an equilibrium analysis and a discussion of the results. Section 4 extends the model by endogenizing officers' post assignments. Finally, the paper concludes with a discussion of the potential implications of the theory.

## 2 Does Class Influence the Likelihood of Police Violence?

Since the police are arguably the state agents with whom citizens interact the most (Lerman and Weaver 2014b), the issue of policing disparities becomes a central concern of democratic citizenship (González 2017; Magaloni and Rodriguez 2020). As such, several studies have been undertaken to examine police-citizen interactions across different class contexts.

For instance, Daniel Brinks (2007) shows how in Uruguay, Brazil, Argentina, and El Salvador, people from low socioeconomic status are more likely to be victims of arbitrary police killings, and how, in Brazil and Argentina, low SES also shapes the judicial outcomes of these killings. When examining other outcomes, socioeconomic status might also affect citizens' interactions with the police. For example, using experimental data collected in two independent field experiments in Mexico and Malawi, Fried, Lagunes, and Venkataramani (2010) and Robinson and Seim (2018), respectively, find that traffic police are more likely to request bribes from drivers signaling low socioeconomic status. The authors suggest that the differential treatment derives from officers' concerns that noticeable high-income drivers would exact retribution. As Yanilda González (2017) argues, these disparities create a type of *stratified citizenship*, characterized by an uneven state response to its citizens.

While these studies provide insights into how citizens' socioeconomic status influences their likelihood of being victimized by the police, they are inadequate for comprehensively examining police violence. The experimental methodologies utilized in Fried, Lagunes, and Venkataramani (2010) and Robinson and Seim (2018) to investigate the correlation between socioeconomic status and the propensity of officers to solicit bribes are ethically questionable when applied to understanding the roots of police violence. The work of Brinks (2007) relies on administrative data to estimate class discrimination in policing. Administrative data may contain incomplete information, leading

to biases in the estimation of the conditional probability of being attacked by the police given class, especially if the police discriminate on the basis of class when they decide with whom they interact. However, if anything, if the police do stop citizens based on class, these effects would be underestimated, implying that Brink's argument would still hold.

Overall, the use of administrative data to study disparities in policing tactics has at least two limitations: 1) the data on police violence incidents is often incomplete. Not all cases of police violence get reported, and interactions that are free of violence tend not to get monitored. 2) We cannot observe all the encounters that did not occur. That is, those citizens that the police could have stopped but restrained from doing. These limitations suggest challenges in using empirical evidence to analyze class disparities in policing.

Despite their limitations, the studies presented above offer valuable insights into the dynamics of class in citizen-police interactions. They provide empirical evidence suggesting a potential explanation for police violence: disparities in access to political institutions render citizens from lower socioeconomic backgrounds more susceptible to brutality by law enforcement. This is because police officers may perceive a lower risk of retribution from wealthier citizens, leading them to exercise restraint when interacting with individuals from affluent backgrounds. Drawing from this literature, I propose a model in the subsequent section and present its equilibrium analysis for the baseline setup game.

# 3 A Game Theoretical Approach to Analyzing Class Disparities in Police Violence

My model examines how citizens' uneven access to the judicial system (i.e., class) might shape police-citizen encounters. I employ these disparities in access to judicial institutions to model class. However, this parameter can also capture differences across other dimensions, such as rural-urban divides or across ethnic groups. The model assumes that police officers only care about class insofar as it affects citizens' ability to lodge a complaint successfully. While this assumption might not hold in reality as there may be other reasons why the police might discriminate based on class, it allows me to examine how institutional biases of a state institution might be passed on to another.

The experimental studies mentioned above shed some light on the mechanisms through which class might operate, leading to unfair treatment by the police. They indicate that the police abstain from targeting citizens from high SES because they fear retribution. Wealthy citizens might have the political connections and the economic capacity to influence their livelihoods. While the underlying reasons why police officers target poor citizens in corruption certainly differ from why they engage in violence towards civilians, I propose the same mechanism—fear of retribution—for the case of police violence. That is, police officers are more likely to abstain from using excessive force when they interact with wealthy citizens because they would expect retribution.

The logic behind this model is that wealthy citizens can afford the cost of hiring lawyers and have the time and resources to expect their complaints to be more effective. Additionally, it might be more challenging for poor citizens to bear the cost of denouncing, and they might have little faith that their complaints will be heard. Needless to say, there are other potential macro-level explanations for police abuses. For instance, the police might perceive the poor citizen as more likely to commit a crime. Based on these perceptions, they might be more inclined to use violence when interacting with the poor.

**Institutional Discrimination and Structural Accounts of Disparities in Policing**

The mechanism briefly outlined above builds on theories of institutional discrimination (Small and Pager 2020), which highlight how policies or institutions that appear neutral on their face can nonetheless generate systematic disadvantages for certain groups, not because of intentional bias, but because of the structural conditions in which they operate.

For instance, Small and Pager (2020) showcase how referral-based hiring can systematically disadvantage racial minorities. Because social networks tend to be racially homogenous, employees in majority-white firms are more likely to refer other white candidates, thereby perpetuating the racial composition of the organization (Arrow 1998.) Crucially, this dynamic does not depend on the referring employees holding prejudiced views (as in taste-based discrimination) or making assumptions about group traits (as in statistical discrimination). Instead, the exclusionary outcome emerges from the interaction between an apparently neutral practice and a structurally unequal social landscape.

Similarly, in the case of police violence, class-based disparities can arise even without officers holding prejudiced views about poorer citizens or making assumptions about their behavior. Although not the focus of this paper, one form of structural discrimination in policing can emerge through the use of place-based strategies, such as "hot spot" policing. Because hot spots are assigned based on reported crime rates, this approach tends to concentrate police presence in neighborhoods with historically higher crime—often low-income areas (Weisburd 2016). Thus, even if the likelihood of aggression in any individual encounter is equal across neighborhoods, the increased number of encounters in poorer areas might raise residents' overall exposure to police violence.

The model introduced below builds on and formalizes the logic underlying institutional discrimination. Specifically, it captures how seemingly neutral accountability practices in the judicial system, such as complaint mechanisms, can permeate interactions between citizens and police, producing unequal outcomes that disproportionately affect the poor.

A complaint mechanism is intended to provide access to justice; however, it often requires resources that are unevenly distributed across social groups. Filing a complaint typically demands time to travel to court, familiarity with procedural requirements, and access to legal representation—all of which are more accessible to wealthier citizens. Even when poorer citizens manage to file complaints, the likelihood that their claims are successfully resolved may depend on their ability to afford better legal advocacy.

This structural disparity is embedded in the model by assuming that poorer citizens face higher costs when filing complaints about police misconduct and have a lower likelihood of success. Because the legal system imposes greater burdens on those with fewer resources, these citizens are less likely to pursue retribution. Officers, in turn, choose whether to engage in aggressive behavior based on their policing style. Those who favor aggressive tactics understand that poorer residents are less likely to respond with formal complaints. As a result, aggressive officers are more likely to self-select into neighborhoods where residents face higher costs to seeking justice. The next section introduces and motivates the model's assumptions about officers' preferences for aggressive versus non-aggressive tactics.

**Preferences for Police Tactics**

Because violent policing is not always visible, misconduct can be difficult to monitor—even when police departments attempt to hold abusive officers accountable. Just as principal-agent theory would predict, the police departments—the principal—are unable to fully monitor the behavior of the police officers—the agent. As a result, police officers have substantial discretion in the way they enforce the law, or as Lipsky (1980) would put it, "policemen [are street-level bureaucrats who] decide who to arrest and whose behavior to overlook." These issues are even starker in developing countries that do not have strong institutions for internal and external police monitoring or accountability.

Thus, since the police enjoy a degree of freedom to decide how and on whom they enforce the law, individual traits of police officers such as aggressiveness or agreeableness, and their assessment of the risk in a citizen-police encounter become relevant for policing; Police officers might deem some tactics more appropriate than others to carry out their duties (Brown 1981; Clark 2024). The model incorporates the importance of these individual traits by allowing police officers to vary in their preferences for different tactics—aggressive or non-aggressive. Aggressive tactics can involve (e.g.) the use of excessive force, assault, or battery. In contrast, non-aggressive tactics are two-fold: police need to promote collaboration, and citizens need to decide to cooperate. Citizens' cooperation typically entails providing information to the police or not resisting police stops. I discuss the role of cooperation in the next section.

For simplicity, the model only considers the actions that police officers can choose when interacting with citizens. Naturally, the use of aggressive policing tactics often transcends an agent's decision to engage in violence. Torture was an institutionalized practice of the military dictatorships in Argentina, Uruguay, Brazil, and Chile. Institutionalized violence by the state was not only present in autocratic regimes in Latin America. State security forces in democratic countries have also systematically employed arbitrary violence against civilians (see, for instance, Magaloni and Rodriguez (2020) and Acemoglu et al. (2020)). For game-theoretical approaches to how institutional factors influence police violence, see Clark (2024) and McCall (2019b).

**Engaging and Cooperating with the Police, and Trust in the Insitution**

An approach to policing highlights the importance of citizen cooperation in police effectiveness. Citizens can inform the police about suspects and collaborate with the police to solve crimes. Scholarship has suggested that citizen cooperation depends on citizens' trust in the police (Sunshine and Tyler 2003; Tyler 2004). If citizens do not trust the police, they are less likely to cooperate with them. The model captures citizens' trust through a parameter representing the share of officers who prefer non-aggressive over aggressive tactics. In the baseline version of the model, citizens are assumed to have accurate beliefs about this share. A more realistic extension introduces an error term to account for potential misperceptions. Additionally, the model assumes that trust is unaffected by incidents of police violence. In practice, however, police violence can significantly influence trust. For example, recent instances of police brutality during the Black Lives Matter protests in the United States have been shown to erode public confidence in the police (Ang et al. 2021).

Trust may also be shaped by factors beyond direct exposure to violence. Recent research in Latin America finds that wealthier citizens are often less trusting of the police than poorer citizens, not because they face more aggression, but because they are more likely to interpret negative encounters—such as corruption or inefficacy—as indicators of institutional failure (Slough and Torreblanca, n.d.). This suggests that trust in the police may itself be a class-based judgment, shaped not only by objective experiences but also by how different social groups interpret those experiences.

In the framework proposed in this paper, poor neighborhoods—characterized by weaker accountability—attract more aggressive officers, which lowers the actual proportion of non-aggressive officers and, accordingly, trust. From this perspective, poorer residents would be expected to trust the police less. However, the model captures only one dimension of trust: the component grounded in structural and behavioral features of police-citizen interactions. By contrast, Slough and Torreblanca (n.d.) highlight how perceptions of trust may diverge from these underlying conditions and instead reflect class-based differences in how citizens interpret negative encounters. Future work could integrate these interpretive processes with structural models of institutional behavior to provide a more comprehensive account of trust in policing.

**Complaint Mechanisms**

My model gives citizens the option to report wrongdoing by the police. This parameter captures the importance of a well-functioning civic judicial system and serves as a proxy for class in the model. In the face of unfair treatment by the police, independent judicial institutions have been proven effective in protecting citizens' rights against abuses by the government and powerful parties in some instances (Hu and Conrad 2020), but not in others. Judicial institutions in high and low-income countries often fail to effectively protect citizens' rights, especially for the marginalized (Méndez, O'donnell, and Pinheiro 1999; Brinks 2007). Moreover, access to judicial institutions is not evenly distributed in developing countries, with significant differences across rural-urban and poor-rich divisions. And even when citizens overcome access limitations, judicial institutions might take too long to solve cases, and in cases of police violence, they often rule in favor of the police (Brinks 2007).

## 3.1 The model

I propose a game-theoretic model of police violence with complete information and two players: a police officer ($O$) and a resident ($R$). Officers are either high ($\theta = h$) or low type ($\theta = l$), determining their preferences for aggressive or non-aggressive police tactics, respectively. Low types prefer to collaborate with citizens, while high types prefer more aggressive techniques. This model feature can reflect individual traits such as aggressiveness, agreeableness, and/or beliefs of what officers consider more effective policing tactics. The resident can choose to engage with the police ($e = 1$), and her decision subsequently determines the probability that the police officer will have the opportunity to resort to aggression, given by $\pi$. More precisely, there is a higher chance that the officer has an opportunity to be aggressive when the resident engages ($\pi_{e=1}$) than when she does not ($\pi_{e=0}$). The intuition behind that assumption is that initiating an interaction with the police residents makes them potentially more vulnerable to the officer's aggression.

Additionally, the model captures class inequalities using two parameters that characterize access to judicial institutions: $\mu$, which represents the likelihood that citizens' complaints about police misconduct are addressed, and $k_r$, which denotes the cost of reporting wrongdoing by the police.

The general assumptions of the model are that police officers differ in their preferences for policing tactics and that citizens' and police' strategies depend on citizens' access to judicial institutions.

Consider the following model setup.

0. Nature randomly chooses a type for the officer. The share of low type officers is $Pr(\theta = l) = q$. $q$ is known by both the officer and the resident.
1. Then, in the first stage of the game, the resident ($R$) decides whether to engage with the officer ($e = 1$) or not ($e = 0$).
2. Next, with probability $\pi(e)$, the officer has an opportunity to use an aggressive tactic ($a = 1$) or $a = 0$ with probability $1 - \pi(e)$.
3. If $a = 1$, the game continues, and the resident chooses whether to report the officer's aggressive behavior ($r = 1$) or not ($r = 0$). Her complaint is effective ($\omega = 1$) with probability $\mu$, and reporting costs $k_r > 0$.

The resident obtains $b_e$ when she engages, and the officer is not aggressive. If the officer chooses to be aggressive, the resident's payoff depends on whether she reports the misconduct and the expected reporting utility. More precisely, if the resident does not complain about the officer's behavior, she obtains $-k_a$ ($k_a > 0$). Under this scenario—when the resident engages and the officer is not aggressive—$b_e$ can be interpreted as the gain in resident security that is only realized when engagement with the police ($e = 1$) is met with non-aggressive tactics ($a = 0$). The logic here is that the resident values an improvement in their security as long as it does not undermine their well-being.

When the resident does report the officer's misconduct, she can potentially receive a higher payoff if her complaint is effective ($\omega = 1$), in which case she receives $b_r - k_r - k_a$. Here, $b_r$ represents compensation for the aggression. When the complaint is not effective, the resident still has to pay the cost of reporting ($k_r$) and the cost of being mistreated by the officer ($k_a$). Formally, the resident's utility is given by:

$$U_R = e(1 - a)b_e - \pi(e)(ak_a(1 - r) + ar(b_r\omega - k_r - k_a)).$$

The officers' utility varies according to their preference for aggressive versus non-aggressive tactics. The high-type officers ($\theta = h$) receive a payoff normalized to zero when they are not aggressive ($a = 0$), regardless of the resident's decision to engage or not. The high-type officer obtains $b_h$ when they choose $a = 1$, and a cost of $p$ would be deducted from this payoff if the resident reports and their complaint is effective. The underlying assumption here is that a high-type officer would typically prefer to perform their duties using aggressive tactics but only abstain from aggression when the threat of punishment, $p\mu$, is high enough. The equation below represents the high-type officer's utility function.

$$U_h = a(b_h - \omega r p)$$

The low-type officer's utility function (see Equation 1) follows a different logic as it depends on whether the resident engages. When the officer chooses not to be aggressive, and the resident does not engage, the officer receives a payoff normalized to zero. However, if the resident engages, the officer obtains $b_l$. This payoff can be interpreted as the benefit police officers receive when a "good" police outcome is realized, such as effectively protecting a citizen or solving a crime. The underlying assumption is that non-aggressive techniques do not yield "good" police outcomes without residents' engagement. If the low-type officer decides to be aggressive, they get a payoff normalized to zero when the resident does not complain or if she complains, but the complaint is not effective. However, if the complaint is effective, the officer incurs a cost of $p$.

$$U_l = e(1 - a)b_l - ar\omega p \tag{1}$$

Note that the low-type officer always prefers to be non-aggressive under all possible combinations of the parameter set ($\mu$, $k_r$, $q$, $b_e$, and $b_r$) (see proofs in the Appendix)x, but their expected utility is higher if the resident engages. Therefore, low-type officers prefer to play a game where the resident is likely to engage.

**Residents' Engagement and non-aggressive Tactics**

This section employs backward induction to delineate the conditions under which officers opt for non-aggressive behavior while the resident ($R$) chooses to engage with the police. First, consider the resident's choice at her last decision node, where she decides whether to report ($r = 1$) or not ($r = 0$). The resident's expected utility is $-k_a$ when $r = 0$ and $b_r\mu - k_r - k_a$ when $r = 1$. That implies that the resident would report if the probability that the complaint is effective, $\mu$, is larger than $\frac{k_r}{b_r}$—the ratio of the cost of reporting over the benefit the resident receives from an effective complaint, $b_r$.

How does the resident's decision to report an aggressive officer affect low and high-type officers' decision to behave aggressively ($a$)? In the case when the resident chooses $r = 0$, the high-type officer obtains a higher payoff if she uses aggressive tactics. More precisely, she receives $b_h$ if $a = 1$ and 0 otherwise. Under the same scenario, the low-type officer gets a payoff of 0 if she chooses $a = 0$ or if the resident does not engage; If the resident engages, the low-type officer, $O(\theta = l)$, gets $b_l$. From that follows that the expected utility for the low type officer when $a = 0$ is equal or larger than her expected utility when $a = 1$. That implies that both officers would choose their preferred tactic when the resident does not report. That is, the low-type officer, $O(\theta = l)$, chooses $a = 0$ and gets $b_l$, and the high-type officer, $O(\theta = h)$, chooses $a = 1$ and gets $b_h$.

Now, consider the case when the resident reports misconduct (i.e., $\mu \geq \frac{k_r}{b_r}$.) The high-type officer's expected utility is $b_h - p\mu$ if they choose $a = 1$ and 0 if $a = 0$. From that follows that the high-type officer would prefer to be non-aggressive if and only if $\mu \geq \frac{p}{b_h}$. Here, the low-type officer gets $-p\mu$ if she chooses $a = 1$, and $b_l$ if $a = 0$, which means the low-type officer's expected utility when $a = 0$ is greater or equal than her expected utility when $a = 1$ if the resident reports.

Now, if it is not sequentially rational to report police misconduct and thus the resident does not report and therefore both officers select their preferred action; That is, the low-type officer ($O(\theta = l)$) chooses $a = 0$, and the high type ($O(\theta = h)$) chooses $a = 1$, the expected utility of not engaging and engaging are respectively given by:

$$EU(e=0, r^*=0) = q0 + (1-q)(-k_a)\pi_0, \text{ and}$$
$$EU(e=1, r^*=0) = qb_e + (1-q)(-k_a)\pi_1.$$

Therefore the resident engages when

$$\begin{aligned} qb_e + (1-q)(-k_a)\pi_1 &\geq (1-q)(-k_a)\pi_0 \\ qb_e &\geq (1-q)(k_a)(\pi_1 - \pi_0) \\ q &\geq \frac{k_a d}{b_e + k_a d}, \end{aligned} \tag{2}$$

where $d = \pi_1 - \pi_0$. That means that when the cost that residents pay if the officer is aggressive is large, the share of low-type officers has to be close to 1 for residents to want to engage with the police, holding $b_e$ at low levels. In contrast, when $k_a$ is low, residents would almost always prefer to engage with the police.

Now, if it is sequentially rational to report aggression by the police, the resident would prefer to engage with the police if and only if

$$\begin{aligned} EU(e=1, r^*=1) &\geq EU(e=0, r^*=1) \\ qb_e + (1-q)(b_r\mu - k_a - k_r)\pi(c=1) &\geq (1-q)(b_r\mu - k_a - k_r)\pi_0 \\ qb_e &\geq (1-q)(k_a + k_r - b_r\mu)(\pi_1 - \pi_0). \end{aligned}$$

Which means that the probability of interacting with a low-type officer, $q$, has to satisfy the following condition to ensure residents' engagement with the police.

$$q \geq \frac{(k_a + k_r - b_r\mu)d}{(b_e + (k_a + k_r - b_r\mu)d)} \tag{3}$$

Similarly to the case of no reporting, if the costs of aggression ($k_a$) and reporting ($k_r$) are very high, and the benefit received by the resident from a positive interaction with the police officer, $b_e$, is rather small, the proportion of low-high type officer has to be also high for residents to prefer to engage with the police.

**Proposition 3.1** (The resident does not report misconduct)**.** *If it is not optimal for the resident to report misconduct (i.e., $\mu \leq \frac{k_r}{b_r}$), residents will engage with the police if $q \geq \frac{k_a d}{b_e + k_a d}$. In such a case, the high-type officer will always choose to be aggressive.*

**Proposition 3.2** (The resident reports misconduct)**.** *If the resident reports misconduct (i.e., $\mu \geq \frac{k_r}{b_r}$), residents will engage with the police if $q \geq \frac{(k_a + k_r - b_r\mu)d}{(b_e + (k_a + k_r - b_r\mu)d)}$. In such a case, the high-type officer will be non-aggressive if it satisfies*

$$\mu \geq \frac{p}{b_h}.$$

The two propositions above indicate that a sufficiently large probability of a successful complaint induces residents' engagement with the police and judicial institutions and positive behavior from aggressive officers. Assuming that residents report, the high-type officer's expected utility is zero if the condition that ensures non-aggression by this officer is satisfied, and $b_h - p\mu$ when it is not satisfied. This implies that high-type officers would prefer to work in areas (neighborhoods) where the probability of an effective complaint is low, as they can be aggressive without the threat of punishment.

The low-type officers prefer to be non-aggressive under all values of $q$, but they can expect a higher utility when $q$ is large enough ($q > \frac{k_a d}{b_R + k_a d} = \hat{q}_{nr}$) or when residents report. Hence the low-type officer prefers to work in areas where the share of officers of their same type is large enough or where residents can effectively file a complaint. These thresholds highlight how $q$, the proportion of low-type officers, creates an interesting dynamic– by having a large enough share of officers with

preferences for non-aggressive tactics (low types), low-type officers would be more attracted to the job.

Thus, when officers can select the areas where they work– or accept a position at a police department–, low-type officers would populate regions where the share of their type prevails. Residents can expect compensation when they file a complaint, and high-type officers prefer areas with poor judicial institutions. We should expect that officers with preferences for non-aggressive tactics would tend to police wealthier neighborhoods, as characterized by $\mu$ and $k_r$, and police with aggressive tactics would prevail in poor areas conditioning the interaction that residents from different socio-economic backgrounds have with the police. An extension of this model is presented in Section 4 considers a scenario where officers have the autonomy to choose their working areas, or they can accept positions at police departments with specific jurisdiction. In this extended model, officers can decide whether to work in impoverished or affluent neighborhoods.

Proofs are presented in the Appendix.

**Share of officers with preferences for non-aggressive tactics** $q$

The inequalities represented by equations 2 and 3 establish distinct thresholds in the variable $q$, thereby delineating three separate regions within its range. This segmentation is visually depicted in Figure 1, illustrating the resident's interaction with the police in equilibrium. In the graph, the red curve corresponds to scenarios where the resident refrains from reporting incidents of aggression. On the other hand, the black and grey curves signify instances where the resident opts to report, with the grey curve denoting a lower probability ($\mu = 0.3$) of the complaint being acknowledged, and the black curve representing a higher probability ($\mu = 0.9$) of successful complaint resolution. This shows how a diminished probability of a complaint being heard (grey curve) elevates the threshold in $q$ at which residents are inclined to engage with the police, in contrast to scenarios with a larger probability of complaint resolution (black curve). This observation underscores the notion that residents exhibit greater propensity to engage with law enforcement authorities when they perceive a higher likelihood of their complaints being taken seriously and addressed.

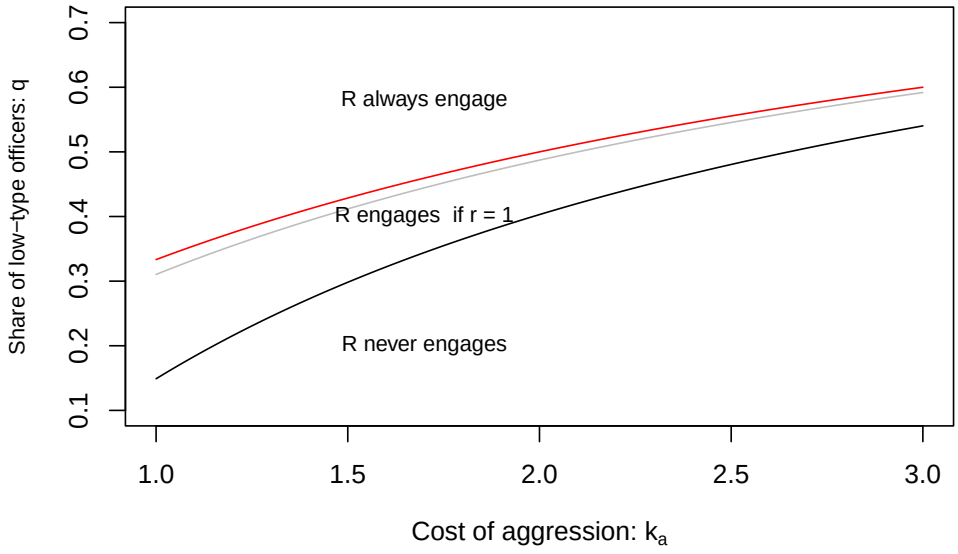


Figure 1: Threshold in q when residents do not report (red), when residents report and the likelihood of an effective complaint is low (grey), when residents report and the likelihood of an effective complaint high (black)

To analyze equilibrium regions, let us first consider the red and black curves ($\mu$ is large) alone. The area above the red line shows the equilibrium in which residents would always engage with the police because $q$ is sufficiently high. The area demarcated by the gap between the red and black curves shows the equilibrium in which residents engage because they can expect a positive utility if they report police misconduct. The gap between the black and red curves illustrates how efficient judicial institutions, characterized by a high $\mu$, can compensate police departments with a low share of officers with preferences for non-aggressive tactics ($q$) to achieve engagement by residents.

Now, let us consider the grey curve ($\mu$ is small). The threshold at which residents engage with the police when they can expect a positive utility of reporting police wrongdoing has increased compared to when $\mu$ is large (black curve.) Thus the region representing the equilibrium in which residents engage if they can expect a positive utility when they report aggression by the officers has shrunk. Therefore in areas characterized by poor judicial institutions– typically poor areas,

as previous research has suggested– the share of officers with preferences for non-aggressive tactics has to be large for residents to prefer to engage with the police.

# 4 Endogenizing Officer Neighborhood Selection

In this section, I present a model where multiple officers can choose to work in one of two different types of neighborhoods: a wealthy ($W$) neighborhood where it is sequentially rational to report police misconduct (i.e., $b_e\mu_W > k_{r,W}$), and an impoverished ($P$) neighborhood where it is not sequentially rational to report misconduct (i.e., $b_e\mu_P < k_{r,P}$). A finite number of officers decide simultaneously where to work between $W$ and $P$ without knowing the choices made by others. Similar to the previous model, the distribution of low-type officers in the entire game is given by $q_0$ and is known.

Once officers select into the two different neighborhoods, residents of each neighborhood learn the value of $q_W$ and $q_P$ in their respective neighborhood. In this way, residents play the same game as in the previous model, meaning that their actions will depend on the values of $q_N$ with $N \in \{W, P\}$.

For the officers, we can imagine that multiple officers are playing the game simultaneously, deciding which neighborhood to patrol. This decision will determine the value of $q_N$– the distribution of low-type officers in the selected neighborhood $N$. The officers receive an additional benefit, $\xi$, when they work in the wealthy neighborhood. This benefit is normally distributed with mean $b_W$ and variance $\sigma^2$. Otherwise, the fixed utility functions remain the same as in the previous model.

Since it is assumed that residents in $P$ do not report police misconduct, the threshold $\hat{q}_P$, at which residents engage with the police, is determined by Inequality 2. On the other hand, since residents in $W$ do report police misconduct, the threshold $\hat{q}_W$, at which residents engage with the police, is given by Inequality 3.

## 4.1 Equilibrium Analysis

When deciding whether to go patrol $P$ or $W$, each low-type officer is uncertain about the actions of other officers, and hence uncertain whether the proportion of officers of their type would be large

enough to yield residents' engagement with the police in either neighborhood.

In contrast, high-type officers face a simpler decision since their utility does not depend on whether residents interact with the police and, as such, is unaffected by the endogenous proportion of officers in each neighborhood, $q_N$. As I show in the appendix, high-type officers will prefer patrolling the impoverished neighborhood $P$ when the expected punishment in the wealthy neighborhood, $\mu_W p$, exceeds the expected benefits of patrolling there, $b_W$. It is easy to see this happening in cases wheres wealthy citizens $W$ have greater political influence and resources, increasing the likelihood ($\mu_W$) of punishment. This can outweigh the benefits of patrolling $W$, which might depend on the perceived safety, social status, or other incentives associated with working in wealthier neighborhoods.

Given the condition $\mu_W p \geq b_W$, under which high-type officers prefer patrolling $P$, I now analyze the strategy of low-type officers and their choice between $P$ and $W$. If all high-type officers choose $P$, then any scenario in which more than one low-type officer patrols $W$ results in $q_W = 1$. In this case, the wealthy neighborhood would be entirely composed of non-aggressive officers, making residents in $W$ feel safe when engaging with the police. Consequently, low-type officers in $W$ receive $b_l$ for using non-aggressive tactics, in addition to $\xi \sim \mathrm{N}(b_W, \sigma^2)$, the stochastic benefit that officers receive from patrolling wealthier neighborhoods.

In contrast, the fixed payoffs for low-type officers in $P$ depend on whether $q_P \geq \hat{q}_P$, where $\hat{q}_P$ is the minimum proportion of officers required for residents to engage with the police in $P$. If this threshold is met and residents engage, low-type officers receive $b_l$. If the threshold is not met and residents do not engage, their optimal fixed payoff in $P$ is zero. Formally, the optimal fixed difference in payoffs, $\Delta U_l^*$, between $P$ and $W$ for low-type officers is expressed as:

$$\Delta U_l^* = \begin{cases} 0 & \text{if } q_P \geq \hat{q}_P, \\ -b_l & \text{otherwise.} \end{cases} \tag{4}$$

It is important to emphasize that $\Delta U_l^*$ represents the difference in fixed utilities and does not vary with the stochastic shock $\xi$.

The threshold defined by equation (4) gives rise to two possible posterior distributions of officers in $P$. Specifically, the proportion of officers in $P$ is given by:

$$q_P = \frac{q_0 \Phi(\frac{\Delta U_l^* - b_W}{\sigma})}{q_0 \Phi(\frac{\Delta U_l^* - b_W}{\sigma}) + (1 - q_0)},$$

where $q_0$ represents the initial distribution of the proportion of officers in the entire game, and $\Phi(\cdot)$ is the cumulative distribution function of the shock $\xi$, which depends on the two values that $\Delta U_l^*$ can take.

Let $\underline{q}$ and $\overline{q}$ denote the posterior distributions of $q_P$ when $\Delta U_l^* = -b_l$ and $\Delta U_l^* = 0$, respectively. Note that $\underline{q}$ is increasing in $b_l$ (the benefit low-type officers receive when $a = 0, e = 1$), while both $\underline{q}$ and $\overline{q}$ are decreasing in $q_0$.

**Proposition 4.1** (Limited low-type presence in $P$)**.** *There exists a unique equilibrium with a scarcity of low-type officers in $P$, where $q^* = \underline{q}$ . This equilibrium arises under the following condition:*

$$b_e < K_a d \frac{1 - q_0}{q_0 \Phi\left(\frac{-b_W}{\sigma}\right)}. \tag{5}$$

Here, $b_e$ is the benefit residents obtain from a positive police interaction ($e = 1, a = 0$), $k_a$ is the cost residents incur when officers are aggressive toward them, and $d$ represents the difference between the likelihood of an opportunity for aggression when residents engage versus when they do not engage with the police. It is assumed that $d > 0$.

- **Implication** (Class-based policing disparities) The Limited low-type presence in $P$ equilibrium provides a framework to explain class-based policing disparities.

As characterized in Proposition 4.1, given the incentives for residents of a positive interaction with the police ($b_e$) and the disincentives of a negative interaction with the police ($k_a$), the low proportion of low-type officers in $P$ ($q_P = \underline{q}$) leads to systemic differences in resident-police interactions across neighborhoods.

More precisely, in $P$, the scarcity of low-type officers entails that residents are more likely to interact with high-type (aggressive) officers, resulting in more frequent abusive encounters. Meanwhile in $W$, under this equilibrium[2], the presence of only non-aggressive officers ensures that residents

[2]There also exists a mixed-strategy equilibrium in which not all high-type officers patrol $P$, although a higher

consistently have positive experiences with the police.

Notably, disparities can also arise under two additional conditions. First, when $b_W$ (the expected benefit of patrolling the wealthy neighborhood) is positive and the stochastic environment is predictable (small $\sigma$), the likelihood of low-type officers patrolling $P$ diminishes further. In this context, the combination of a clear, attractive benefit and low uncertainty associated with patrolling $W$ makes it a more appealing choice for low-type officers, reducing their willingness to patrol $P$. Second, as the initial proportion of low-type officers in the game ($q_0$) decreases, the right-hand side of Inequality 5 grows, making it easier for any value of $b_e$ to satisfy the condition.

Together, these conditions give rise to class-based disparities in policing, driven by asymmetries in access to judicial institutions. These asymmetries establish a self-reinforcing dynamic wherein the neighborhood selection of non-aggressive officers and residents' engagement with the police are mutually dependent. Residents are more likely to engage with the police when they expect interactions with non-aggressive officers, and non-aggressive officers, in turn, are drawn to neighborhoods where they can anticipate positive engagement from residents.

In neighborhoods with limited access to judicial institutions, residents are often reluctant to report misconduct due to factors such as the time required to file complaints, captured by $k_{r,P}$, or the inability to navigate complex systems or afford legal representation, captured by $\mu_P$. This reluctance creates an environment that attracts aggressive officers, as weak accountability mechanisms and low deterrence against misconduct prevail. In contrast, in neighborhoods with stronger institutional access, where residents expect their complaints to be addressed, aggressive officers are discouraged by the higher likelihood of facing consequences for their actions.

This improved accountability fosters a cycle of positive interactions between police and residents. For non-aggressive officers, such neighborhoods become more appealing because they anticipate greater resident engagement, which they believe supports better policing outcomes.

For simplicity, this equilibrium analysis assumes it is not sequentially rational for residents in $P$ to report misconduct. However, the findings remain even when it is sequentially rational for residents in $P$ to report misconduct, provided that $\mu_W > \mu_P$.

proportion still choose $P$ over $W$. In this equilibrium, high-type officers adopt a mixed strategy, dividing their patrols between $P$ and $W$. Such equilibrium is characterized by parameters $b_W$, $\sigma$, $p$, and $\mu_W$.

**Proposition 4.2** (High presence of low-type officers in $P$.)**.** *There exists a unique equilibrium where the proportion of low-type officers in the impoverished neighborhood ($P$) is high, given by $q^* = \overline{q}$. This equilibrium occurs when residents derive substantial benefits from positive interactions with the police ($b_e$), such that the following inequality is satisfied:*

$$b_e > K_a d \frac{1 - q_0}{q_0 \Phi\left(\frac{-b_l - b_W}{\sigma}\right)}. \tag{6}$$

In contrast to the equilibrium of limited low-type presence in $P$, where aggressive officers dominate impoverished areas and residents face frequent misconduct, this equilibrium describes a scenario where the proportion of low-type officers in $P$ is high ($q^* = \overline{q}$). When residents derive enough benefits from positive interactions with the police ($b_e$), they are more likely to engage, creating conditions that attract non-aggressive officers.

$K_a d$ represents the baseline deterrent effect of aggressive behavior. $K_a$ is the cost to residents when officers are aggressive, and $d$ captures the increased likelihood of aggression when residents engage with the police. Together, they quantify how much aggression discourages engagement.

The factor $\frac{1-q_0}{q_0 \Phi\left(\frac{-b_l - b_W}{\sigma}\right)}$ magnifies the baseline deterrent effect ($K_a d$) based on the prior proportion of non-aggressive officers ($q_0$) and the relative attractiveness of patrolling $W$. When $q_0$ is small or $W$ is highly appealing to officers, the deterrent effect is amplified, requiring a significantly higher $b_e$ to achieve an equilibrium where $P$ has a high presence of low-type officers.

**Proposition 4.3** (Multiple equilibria)**.** *Multiple equilibria arise when:*

$$K_a d \frac{1 - q_0}{\Phi\left(\frac{-b_W}{\sigma}\right)} \leq b_e \leq K_a d \frac{1 - q_0}{\Phi\left(\frac{-b_l - b_W}{\sigma}\right)}$$

These inequalities are satisfied when $b_l$ is non-negative.

Proposition 4.3 identifies conditions under which the model produces multiple equilibria. Specifically, when $b_e$ (the benefit residents receive from positive police interactions) falls within the given range, both the low-type presence in $P$ ($\underline{q}$) and high-type presence in $P$ ($\overline{q}$) are possible outcomes.

Figure 2 illustrates the equilibria described in Propositions 4.1 to 4.3. The vertical axis represents

the posterior distribution of low-type officers ($q_P$), while the horizontal axis shows possible values of $q_P$. Each grid in the plot corresponds to a different value of $b_e$, resulting in a distinct engagement threshold $\hat{q}_P$. The lower plotted tick line represents $\underline{q}$, while the upper tick line corresponds to $\overline{q}$. The jump between these two lines occurs at $\hat{q}_P$, indicated by the dashed horizontal line. Additionally, the diagonal line $x = y$ represents the posterior distribution in equilibrium ($q^*$).

These plots show how changes in $b_e$ affect the distribution of low-type officers. As $b_e$ increases, engagement thresholds ($\hat{q}_r$) decline, making residents more likely to engage with the police, atracting a higher proportion of non-aggressive officers. The plots in the top row depict cases where a unique equilibrium arises, as characterized by Proposition 4.1, with low $b_e$ resulting in a scarcity of low-type officers in $P$. Intermediate values of $b_e$ satisfy the conditions for multiple equilibria, as outlined in Proposition 4.3. The final plot illustrates scenarios where a high equilibrium emerges, characterized by Proposition 4.2.

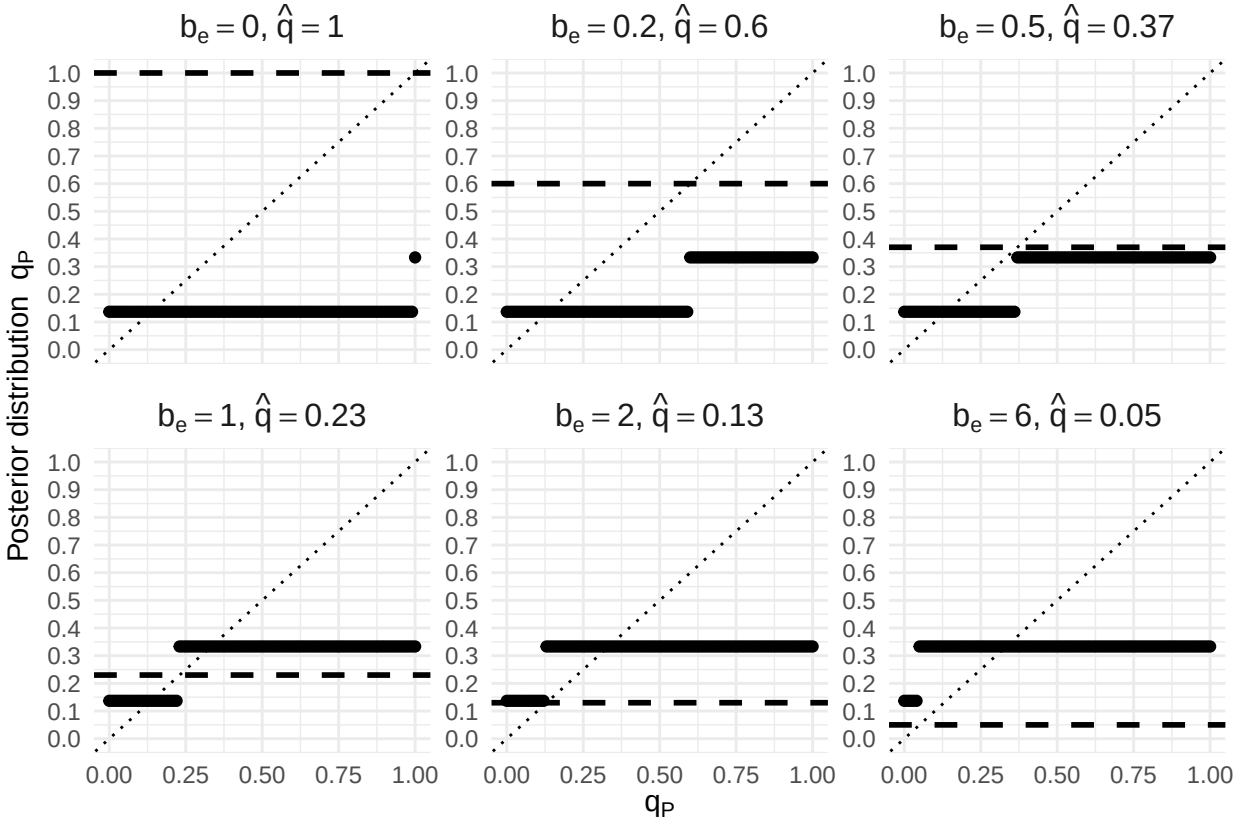


Figure 2: Posterior distribution across different thresholds.

# 5 Conclusion

I introduced a model of police violence with two actors: a police officer and a resident to examine the conditions under which residents cooperate, and officers are not aggressive. Residents' engagement with the police is determined by the parameters defining their access to judicial institutions and the share of officers with preferences for non-aggressive tactics (low-types). More precisely, residents engage with the police if they can expect that the aggressive officer will be punished and where the likelihood of interacting with an aggressive officer is low. These results produce a reinforcing cycle. If officers were to select their posts, high types would prefer areas with poor access to judicial institutions where residents do not report. Low types would be more willing to work in areas where residents can expect their complaints to be heard, and the share of low-type officers is high.

The analysis presented in Section 4 sheds light on the dynamics of officer deployment and its impact on community policing. In scenarios where aggressive officers predominantly populate impoverished neighborhoods with poor access to judicial institutions, while less aggresive officers concentrate in wealthier neighborhoods, a reinforcing cycle emerges. More aggressive officers are drawn to areas where residents are less likely to report misconduct, exacerbating the prevalence of police violence. Conversely, less aggressive officers tend to work in neighborhoods with better access to judicial institutions. This pattern underscores the importance of equitable officer distribution to address disparities in policing practices and promote safer communities.

These findings are in line with recent work by Ba et al. (2021). Their work suggests that through a bidding process, senior officers—typically more experienced and more expensive workers end up in higher-income and predominantly white neighborhoods, and the least experienced officers are sent to areas of high crime where experience is needed the most. They argue that this equilibrium leads to more incidents of police violence, which could be avoided if police departments distributed the assignment of officers more evenly. Future work should consider alternative equilibrium solutions.

# Appendix

### Proofs

**Residents' decision to report**

$$EU(r=1) = b_R\mu - k_r - k_a$$

$$EU(r=0) = -k_a$$

$$EU(r=1) \geq EU(r=0)$$

$$\mu \geq \frac{k_r}{b_R}$$

**Residents' decision to engage when no reporting ($r^* = 0$)**

$$EU(e=0, r^*=0) = q0 + (1-q)(-k_a)\pi(e=0)$$

$$EU(e=1, r^*=0) = qb_e + (1-q)(-k_a)\pi(e=1)$$

$$EU(e=1, r^*=0) \geq EU(e=0, r^*=0)$$

$$qb_R \geq (1-q)(k_a)(\pi(e=1) - \pi(e=0))$$

Let $d = \pi(e=1) - \pi(e=0)$

$$q(b_R + k_a d) \geq k_a d$$

$$q \geq \frac{k_a d}{(b_e + k_a d)}$$

**Residents' decision to engage when reporting ($r^* = 1$)**

$$EU(e=0, r^*=1) = q0 + (1-q)(b_R\mu - k_a - k_r)\pi(e=0)$$

$$EU(e=1, r^*=1) = qb_R + (1-q)(b_R\mu - k_a - k_r)\pi(e=1)$$

$$EU(e=1, r^*=1) \geq EU(e=0, r^*=1)$$

$$qb_R \geq (1-q)(k_a + k_r - b_R\mu)(\pi(e=1) - \pi(e=0))$$

Let $d = \pi(e=1) - \pi(e=0)$

$$q(b_R + (k_a + k_r - b_R\mu)d) \geq (k_a + k_r - b_R\mu)d$$

$$q \geq \frac{(k_a + k_r - b_R\mu)d}{(b_R + (k_a + k_r - b_R\mu)d)}$$

**High-type officers' decision not to be aggressive when $r^* = 1$**

$$EU(a = 0, \theta = h) = 0$$

$$EU(a = 1, \theta = h) = b_h - p\mu$$

$$EU(a = 0, \theta = h) \geq EU(a = 1, \theta = h)$$

$$\mu \geq \frac{p}{b_h}$$

**Low-type officers' decision not to be aggressive when $r^* = 1$**

$$EU(a = 1, \theta = l) = -p\mu$$

$$EU(a = 0, \theta = l) = b_l$$

$$EU(a = 0, \theta = l) \geq EU(a = 1, \theta = l)$$

$$mu \leq -\frac{b_l}{p}$$

**High-type officers' decision not to be aggressive when $r^* = 0$**

$$EU(a = 0, \theta = h) = 0$$

$$EU(a = 1, \theta = h) = b_h$$

$$EU(a = 0, \theta = h) \geq EU(a = 1)$$

$$b_h \leq 0$$

**Low-type officers' decision to be aggressive when $r^* = 0$**

$$EU(a = 1, \theta = l) = 0$$
$$EU(a = 0, \theta = l) = b_l$$
$$b_l \geq 0$$

**Proof of proposition 4.1**

**Proof: High-type officers prefer $P$** The expected utility for high-type officers in $P$ and $W$ is given by:

$$E[U_h(P)] = b_a \quad \text{and} \quad E[U_h(W)] = b_a - \mu_W p + b_W,$$

respectively. High-type officers will prefer $P$ when:

$$\mu_W p \geq b_W.$$

**Proof: Strategy of low-type officers when high-type officers prefer $P$**

When the threshold of low-type officers required for residents to engage with the police in $P$, $\hat{q}_P$, is greater than both $\underline{q}$ and $\overline{q}$, there is a unique equilibrium where $q_P^* = \underline{q}$.

If $\hat{q}_P > \overline{q}$, it implies $\hat{q}_P > \underline{q}$ as well. This ensures that any posterior distribution for $q_P$ satisfies $q_P < \hat{q}_P$, leading to:

$$q_P^* = \underline{q} = \frac{q_0 \Phi\left(\frac{-b_l - b_W}{\sigma}\right)}{q_0 \Phi\left(\frac{-b_l - b_W}{\sigma}\right) + (1 - q_0)}.$$

To determine the condition for this equilibrium, the following inequality must hold:

$$\hat{q}_P > \overline{q}.$$

Substituting the expression for $\hat{q}_P$:

$$\frac{k_a d}{b_e + k_a d} > \overline{q}.$$

Rearranging:

$$b_e < \frac{k_a d(1-\overline{q})}{\overline{q}}.$$

Substituting $\overline{q} = \frac{q_0\Phi\left(\frac{-b_W}{\sigma}\right)}{q_0\Phi\left(\frac{-b_W}{\sigma}\right)+(1-q_0)}$, we have:

$$b_e < k_a d \frac{\left(1 - \frac{q_0\Phi\left(\frac{-b_W}{\sigma}\right)}{q_0\Phi\left(\frac{-b_W}{\sigma}\right)+(1-q_0)}\right)}{\frac{q_0\Phi\left(\frac{-b_W}{\sigma}\right)}{q_0\Phi\left(\frac{-b_W}{\sigma}\right)+(1-q_0)}}.$$

Simplifying further:

$$b_e < \frac{k_a d(1-q_0)}{q_0\Phi\left(\frac{-b_W}{\sigma}\right)}. \quad \blacksquare$$

**Proof of proposition 4.2**

This proof work in a similar way as the proof of Proposition 4.1 When the threshold of low-type officers required for residents to engage with the police in $P$, $\hat{q}_P$, is smller than both $\underline{q}$ and $\overline{q}$, there is a unique equilibrium where $q_P^* = \overline{q}$.

To determine the condition for this equilibrium, the following inequality must hold:

$$\hat{q}_P < \underline{q}.$$

Substituting the expression for $\hat{q}_P$:

$$\frac{k_a d}{b_e + k_a d} < \underline{q}.$$

Rearranging:

$$b_e > \frac{k_a d(1-\underline{q})}{\underline{q}}.$$

Substituting $\underline{q} = \frac{q_0\Phi\left(\frac{-b_l-b_W}{\sigma}\right)}{q_0\Phi\left(\frac{-b_l-b_W}{\sigma}\right)+(1-q_0)}$, we have:

$$b_e > k_a d \frac{\left(1 - \frac{q_0\Phi\left(\frac{-b_W}{\sigma}\right)}{q_0\Phi\left(\frac{-b_l-b_W}{\sigma}\right)+(1-q_0)}\right)}{\frac{q_0\Phi\left(\frac{-b_l-b_W}{\sigma}\right)}{q_0\Phi\left(\frac{-b_W}{\sigma}\right)+(1-q_0)}}.$$

Simplifying further:

$$b_e > \frac{k_a d(1-q_0)}{q_0\Phi\left(\frac{-b_l-b_W}{\sigma}\right)}. \quad \blacksquare$$

## References


Acemoglu, Daron, Leopoldo Fergusson, James Robinson, Dario Romero, and Juan F. Vargas. 2020. "The Perils of High-Powered Incentives: Evidence from Colombia's False Positives." *American Economic Journal: Economic Policy* 12 (3): 1–43. https://doi.org/10.1257/pol.20180168.

Ang, Desmond, Panka Bencsik, Jesse Bruhn, and Ellora Derenoncourt. 2021. "Police Violence Reduces Civilian Cooperation and Engagement with Law Enforcement." *HKS Working Paper No. RWP21-022*. https://ssrn.com/abstract=3920493 .

Arrow, Kenneth J. 1998. "What Has Economics to Say About Racial Discrimination?" *Journal of Economic Perspectives* 12 (2): 91–100. https://doi.org/10.1257/jep.12.2.91.

Ba, Bocar, Patrick Bayer, Nayoung Rim, Roman Rivera, and Modibo Sidibe. 2021. "Police Officer Assignment and Neighborhood Crime." *NBER Working Paper No. W29243*. https://ssrn.com/abstract=3922517 .

Brinks, Daniel M. 2007. *The Judicial Response to Police Killings in Latin America: Inequality and the Rule of Law*. Cambridge University Press. https://doi.org/10.1017/CBO9780511551130.

Brown, Michael K. 1981. *Working the Street Police Discretion and the Dilemmas of Reform / Michael k. Brown.* Publications of Russell Sage Foundation. New York: Russell Sage Foundation.

CEDHJ. 2020. "Informe Especial Sobre El Caso de Alejandro Giovanni Lopez Ramirez Del Municipiio de Ixtlahuacan de Los Membrillos." http://historico.cedhj.org.mx/recomendaciones/inf.%20especiales/2020/INFORME%20ESPECIAL%20SOBRE%20EL%20CASO%20GIOVANNI.pdf.

Clark, Tom S. 2024. "A Political-Economic Model of Police Administration." *The Journal of Politics*, October. https://doi.org/10.1086/729933.

Clark, Tom S., Elisha Cohen, Adam N. Glynn, Michael Leo Owens, Anna Gunderson, and Kaylyn Jackson Schiff. 2023. "Are Police Racially Biased in the Decision to Shoot?" *The Journal of Politics* 85 (3): 826–42. https://doi.org/10.1086/723973.

Fried, Brian, Paul Lagunes, and Atheendar Venkataramani. 2010. "Corruption and Inequality at the Crossroad: A Multi-Method Study of Bribery and Discrimination in Latin America." *Latin American Research Review* 45 (November). https://doi.org/10.1353/lar.0.0107.

Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. 2007. "An Analysis of the New York City Police Department's 'Stop-and-Frisk' Policy in the Context of Claims of Racial Bias." *Journal of the American Statistical Association* 102 (479): 813–23. https://doi.org/10.1198/016214506000001040.

González, Yanilda María. 2017. "'What Citizens Can See of the State': Police and the Construction of Democratic Citizenship in Latin America." *Theoretical Criminology* 21 (4): 494–511. https://doi.org/10.1177/1362480617724826.

Hu, Shengkuo, and Courtenay R Conrad. 2020. "Monitoring via the Courts: Judicial Oversight and Police Violence in India." *International Studies Quarterly* 64 (3): 699–709. https://doi.org/10.1093/isq/sqaa039.

Knox, Dean, Will Lowe, and Jonathan Mummolo. 2020. "Administrative Records Mask Racially Biased Policing." *American Political Science Review* 114 (3): 619–37. https://doi.org/10.1017/S0003055420000039.

Lerman, Amy E., and Vesla M. Weaver. 2014a. "Staying Out of Sight? Concentrated Policing and Local Political Action." *The Annals of the American Academy of Political and Social Science* 651: 202–19. http://www.jstor.org/stable/24541702.

———. 2014b. "Staying Out of Sight? Concentrated Policing and Local Political Action." *The Annals of the American Academy of Political and Social Science* 651: 202–19. http://www.jstor.org/stable/24541702.

Lipsky, Michael. 1980. "Street Level Bureaucracy: Dilemmas of the Individual in Public Services." In, 3–12. Russell Sage Foundation.

Magaloni, Beatriz, and Luis Rodriguez. 2020. "Institutionalized Police Brutality: Torture, the Militarization of Security, and the Reform of Inquisitorial Criminal Justice in Mexico." *American Political Science Review* 114 (4): 1013–34. https://doi.org/10.1017/S0003055420000520.

McCall, Andrew. 2019a. "Resident Assistance, Police Chief Learning, and the Persistence of Aggressive Policing Tactics in Black Neighborhoods." *The Journal of Politics* 81 (3): 1133–42. https://doi.org/10.1086/703385.

———. 2019b. "Resident Assistance, Police Chief Learning, and the Persistence of Aggressive Policing Tactics in Black Neighborhoods." *The Journal of Politics* 81 (3): 1133–42. https://doi.org/10.1086/703385.

Méndez, Juan E., Guillermo O'donnell, and Paulo S. Pinheiro. 1999. "The (Un)rule of Law and the Underprivileged in Latin America." In.

O'Flaherty, Brendan, and Rajiv Sethi. 2008. "Racial Stereotypes and Robbery." *Journal of Economic Behavior & Organization* 68 (February): 511–24. https://doi.org/10.1016/j.jebo.2008.06.007.

Peeples, Lynne. 2020. "What the Data Say about Police Brutality and Racial Bias — and Which Reforms Might Work." *Nature* 583 (7814): 22–24. https://doi.org/10.1038/d41586-020-01846-z.

Robinson, Amanda Lea, and Brigitte Seim. 2018. "Who Is Targeted in Corruption? Disentangling the Effects of Wealth and Power on Exposure to Bribery." *Quarterly Journal of Political Science* 13 (3): 313–31. https://doi.org/10.1561/100.00017067.

Seron, Carroll, and Frank Munger. 1996. "Law and Inequality: Race, Gender... and, of Course, Class." *Annual Review of Sociology* 22 (1): 187–212. https://doi.org/10.1146/annurev.soc.22.1.187.

Slough, Tara, and Carolina Torreblanca. n.d. "Class and the Development of Trust in Police in Latin America." *Working Paper*. https://carolina-torreblanca.github.io/files/papers/trust/paper.pdf.

Small, Mario L., and Devah Pager. 2020. "Sociological Perspectives on Racial Discrimination." *Journal of Economic Perspectives* 34 (2): 49–67. https://doi.org/10.1257/jep.34.2.49.

Sunshine, Jason, and Tom R. Tyler. 2003. "The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing." *Law & Society Review* 37 (3): 513–48. https://doi.org/https://doi.org/10.1111/1540-5893.3703002.

Tyler, Tom R. 2004. "Enhancing Police Legitimacy." *The Annals of the American Academy of Political and Social Science* 593: 84–99. http://www.jstor.org/stable/4127668.

Weisburd, David. 2016. "Does Hot Spots Policing Inevitably Lead to Unfair and Abusive Police Practices, or Can We Maximize Both Fairness and Effectiveness in the New Proactive Policing?" *University of Chicago Legal Forum* 2016 (16). https://chicagounbound.uchicago.edu/uclf/vol2016/iss1/16/?utm_source=chicagounbound.uchicago.edu%2Fuclf%2Fvol2016%2Fiss1%2F16&utm_medium=PDF&utm_campaign=PDFCoverPages.