

# Smartphone Data Reveal Neighborhood-Level Racial Disparities in Police Presence

M. Keith Chen, Katherine L. Christensen, Elicia John,  
Emily Owens, Yilin Zhuo\*

February 16, 2023

## Abstract

Research on policing has focused on documented actions such as stops and arrests—less is known about patrols and presence. We map the neighborhood movement of nearly ten thousand officers across 21 of America’s largest cities using anonymized smartphone data. Police spend 0.27% more time in neighborhoods for each percentage point increase in Black residents. This neighborhood-level disparity persists after controlling for density, socioeconomic, and crime-driven demand for policing, and is lower in cities with more Black police supervisors (but not officers). Patterns of police presence statistically explain 55% of the higher arrest rate in more Black neighborhoods.

---

\*Chen: UCLA Anderson School of Management, Los Angeles, CA 90095, USA, keith.chen@anderson.ucla.edu. Christensen: Indiana University, Kelley School of Business, Bloomington, IN 47405, USA, kachris@iu.edu. John: American University, Kogod School of Business, Washington, DC 20016, USA, eliciaj@american.edu. Owens: Department of Criminology, Law & Society, Department of Economics, University of California Irvine, Irvine, CA 92697, USA, egowens@uci.edu. Zhuo: UCLA Anderson School of Management, Los Angeles, CA 90095, USA, yilin.zhuo.phd@anderson.ucla.edu. Acknowledgements: We thank Milind Rajavasireddy for excellent research assistance, as well as Bocar Ba, Kitt Carpenter, Amanda Geller, Michael Norton, Broderick Turner, and seminar participants at the Harvard Kennedy School, UCLA Anderson, UCI, USC, the University of Michigan, the University of Virginia, and the Virtual Crime Economics (ViCE) Seminar for helpful comments.

According to FBI statistics, Black people in America were arrested at twice the rate of White people in 2019 ([OJJDP Statistical Briefing Book 2019](#)). A large literature explores the causes of racial disparities in police enforcement actions, such as stops, searches, and arrests, including differences in socioeconomic status, criminal activity, and biased decision making by police officers ([Banaji et al. 2021](#); [Banks et al. 2006](#); [Hoekstra and Sloan 2022](#); [Rucker and Richeson 2021](#)). Disparities in police enforcement are highly consequential for impacted civilians, but may not fully reflect disparities in the entirety of what it means for a person, or an area, to be “policed.”

In this paper, we provide evidence on the following question: do police departments differentially patrol the more heavily Black, Hispanic, or Asian neighborhoods in their cities? A priori, police presence can either help, or harm, communities. Police presence can deter crime. It can also influence when and where crime is officially recorded. Finally, police presence is necessary for a stop, search or arrest to occur. Thus, detailed information on the neighborhoods where officers work during their shifts and on how the racial composition of neighborhoods varies both across and within cities can identify sources of disparities in later criminal justice outcomes. Unfortunately, few departments record detailed data on where officers are during their shifts, and even fewer make it available to researchers in a standardized way.

We use anonymized smartphone location data to identify and track the movements of individual police officers on patrol in 21 of the largest cities in the United States. We measure police presence as the total number of officer-hours spent in a census block group (a “neighborhood” with roughly 1,000 residents) over a ten-month period (Feb 2017 - Nov 2017), when the officer is moving through a neighborhood at 50 mph or less. These data identify where police spend their time, and allow us to evaluate spatial patterns of policing at scale while protecting officer privacy.

Using these data, we quantify how patterns of socioeconomic status, crime, social capital, and race relate to local police presence within and across these cities. While we do not

evaluate whether such resource allocation is socially optimal, we document the following facts: (1) Officer presence tends to be higher in non-White neighborhoods, both within and between cities, and there is a large amount of cross-city heterogeneity in this result, (2) Black neighborhoods have the highest officer presence, and though (3) the disparity in officer presence in Asian neighborhoods can be fully explained by neighborhood characteristics, (4) over two thirds of the increased police presence in more Black and Hispanic neighborhoods cannot be explained by neighborhood characteristics.

Generally, geographic analysis of policing at the sub-city level has measured policing in one of two ways. Researchers have studied downstream measures—outcomes of police officer and civilian interactions—and upstream measures—departmental decisions that are made before a civilian interaction (e.g., patrol assignments). Our research builds on a growing literature that examines the role of upstream measures of policing (e.g., in Chicago ([Ba et al., 2021](#)), in Dallas ([Weisburd, 2021](#)), in an English police department ([Vomfell and Stewart, 2021](#)), and in Milan ([Mastrobuoni, 2019](#))). We extend these single city studies in two key ways. First, our smartphone dataset allows us to examine actual police presence in neighborhoods, rather than beat assignment or patrol car locations; this allows us to observe the potentially nontrivial amount of time officers spend outside their assigned patrol locations or their patrol car ([Weisburd, 2021](#)). Second, because our smartphone dataset is independent of city-level decisions to collect or release data ([Goel et al., 2017](#)), we can extend the single-city analyses that typify existing upstream studies to better understand policing within and across 21 of America’s largest cities.

Our neighborhood-level analysis of GPS location data shows that police officers spend more time in places with larger Black, Hispanic, or Asian populations both between and within cities. While controlling for variation in socioeconomic status, social disorganization, and violent crime reduces these disparities, it does not eliminate the disparity in officer time spent in more Black or Hispanic vs. more White neighborhoods. This suggests that social interventions targeted at the “root causes” of crime may be unlikely to eliminate the

racial and ethnic disparities we observe in American policing and confirms qualitative and historical research on upstream police presence across America ([Hinton 2016](#); [Rios 2011](#); [Sharkey 2018](#)), and patterns observed at the city level ([Carmichael and Kent 2014](#)).

While still descriptive, we also explore whether differences in police presence are associated with the racial composition of officers across cities. In contrast with existing single-department studies (e.g., [Hoekstra and Sloan 2022](#); [Ba et al. 2021](#)), our results suggest that the additional police presence in Black neighborhoods is higher in cities where more patrol officers are Black. However, conditional on the share of Black patrol officers, increasing the share of Black front-line supervisors, who direct patrol officer activity, reduces the amount of time spent in Black neighborhoods. While not causal, this highlights the potential role of retention and promotion in police reform aimed at reducing racial disparities in the criminal justice system.

We also provide evidence that the nature of disparities in police presence differs across US cities. While racial disparities in some cities (e.g., Charlotte, NC) are largely associated with spatial differences in socioeconomic status (e.g., income, education, and civic engagement) others persist after controlling for these factors, and for spatial patterns of violent crime (e.g., Austin, TX).

Our work complements existing spatial analyses of downstream measures of policing, which have found that police engage in more enforcement actions in Black neighborhoods ([Geller et al. 2014](#); [Ba et al. 2021](#); [Pierson et al. 2020](#)). In the six cities (including New York City) with publicly available geocoded arrest data, we connect our upstream measures of neighborhood police presence to downstream arrests within that neighborhood. We then separate observed neighborhood arrest disparities into two parts: percent differences in officer-hours spent in a neighborhood and percent differences in arrests per officer-hour spent in that place. We find that differences in where officers spend time explain roughly 55% of the Black-White disparity in neighborhood arrests, conditional on neighborhood characteristics. Officers' higher propensity to make an arrest, conditional on being in a relatively more

Black neighborhood, explains 45% of these disparities.

Taken as a whole, our findings suggest that disparities in exposure to police in the US are associated with both structural socioeconomic disparities and discretionary decision making by police commanders and officers. This study provides novel data on police-civilian interaction to enable additional analyses of the factors driving these observed disparities in hopes of developing policy interventions to mitigate them. Finally, our police presence measure provides a new benchmark against which downstream police actions like stops and arrests may be objectively evaluated.

## 2 Methods

### 2.1 Data

The smartphone location data used in this study were provided by Safegraph and can now be obtained from Veraset, a company that aggregates anonymized location data from a suite of smartphone applications.<sup>1</sup> The smartphone data records “pings” denoting where a specific smartphone is located at a particular point in time. Pings are logged at irregular time intervals, whenever a participating smartphone application requests location information. The modal time between consecutive pings associated with a device is roughly 10 minutes. Our smartphone data covers more than 50 million smartphones, spanning the continental US, in a 10-month period from February 2017 to November 2017. While the dataset contains geolocation information from only a subset of all smartphones, previous studies have found it highly representative of the United States on numerous demographic dimensions (Chen et al. 2021).

We link the smartphone data to two other data sources: 1) police station location data published by the Department of Homeland Security, verified with each city’s open data portal

---

<sup>1</sup>For more information on the Veraset data, see <https://datarade.ai/data-providers/veraset/profile>.

and google maps data, and 2) building rooftop geofence data provided by Microsoft, enabling us to associate each police station’s latitude-longitude location to a geofence that delineates the convex hull of a building’s rooftop boundary. To identify patrol officers in local city neighborhoods, we include police stations categorized as patrol stations, as headquarters, or as unspecified police facilities, resulting in a total of 330 stations across 21 of America’s largest cities. A description of other data sources and data cleaning process can be found in Appendix A. It is important to note the selection of the cities in our sample was based on jurisdictional population and the physical construction of police buildings. Our sample was not determined by the investment the department chose to make in electronic monitoring of officers, or a department’s decision to release the data publicly or enter into a research agreement with external parties (see [Goel et al. \(2017\)](#) for a discussion of these issues in the context of measuring police bias).

## 2.2 Measuring Police Presence

We infer whether a smartphone belongs to an officer by linking smartphone data to police stations’ geofences in several steps. First, if a specific smartphone is observed in a police station geofence at least five days in a month, we identify it as belonging to a police employee in that month. We next infer each smartphone user’s “home” as the smartphone’s modal Geohash-7 (a  $152\text{m} \times 152\text{m}$  grid) that does not include a police station. We identify two home locations, for the early and the latter half of the year, to account for a potential summer move. Then, we identify patrol officers by looking for a specific pattern: leaving home, traveling to a police station, moving around the city (without returning home), returning to the police station, and then going home. The movements of that smartphone between the first and the last station visits are assumed to be the actual locations of a patrol officer while working a “shift.” We require that shifts are bracketed by home visits that are no more than 24 hours apart, and are no shorter than four hours.<sup>2</sup> Under this definition, our officer

---

<sup>2</sup>All results in our analysis are highly robust to limiting our sample to 8 to 12 hour shifts, requiring shifts bracketed by home visits no longer than 18 hours and excluding shifts with long hours spent within

smartphone GPS data sample consists of 9,557 officers that have at least one shift, with a mean shift length of 8.07 hours.<sup>3</sup>

To measure police presence in all census block groups (“neighborhoods”) within the city’s jurisdiction, we look at officers’ smartphone pings when officers are “on shift” and outside of police stations, in a month during which the device appears in a police station on at least 5 days. We conceptualize police presence in a city neighborhood as the number of officer-hours spent in the neighborhood. Specifically, we match police officers’ ping locations to block groups, exclude pings moving faster than 50 mph, and assign the duration of each ping as half of the time between its previous and next ping.<sup>4</sup> We then compute the sum of officer-hours from all officers’ pings observed in the block group across the ten-month period. The resulting estimate of where police spend time on patrol is highly non-uniform, and as our later regression analysis will confirm, is strongly correlated with demographics in ways that produce large racial disparities.

## 2.3 Validity Check

Our study focuses on America’s largest cities. While our data do not capture the universe of police officers in a city, our estimates of the number of officers in a city satisfy many tests of face validity as a measure of police presence. The number of patrol officer devices that we observe across US cities is highly correlated with FBI estimates of police force size ( $\rho = 0.97$  for total count measures,  $\rho = 0.48$  for per capita measures).<sup>5</sup> Further, we can probabilistically impute each device’s “race” using its home census block’s racial composition. There is essentially a one for one unconditional relationship between the imputed racial composition of the police departments in our sample and the racial composition reported by the police stations. These results are available on request.

<sup>3</sup>Figure A.1 in the Online Appendix displays the temporal and spatial pattern of pings for one likely LAPD officer.

<sup>4</sup>Using other constructs of police presence yields qualitatively and quantitatively similar results. Replications of our analysis using the number of distinct officer shifts, alternate (or no) speed thresholds, are available on request.

<sup>5</sup>Appendix Figure A.2 plots the specific values for each city.

the department in the 2016 Law Enforcement Management and Administration Statistics (LEMAS); conditional on the racial composition of the city, a one percentage point increase in our estimate of the percent of the police force that is White (Black, Hispanic, Asian) is associated with a 0.6 (0.6, 0.9, 0.6) percentage point increase in the reported percent of the force that is White (Black, Hispanic, Asian) in the LEMAS. <sup>6</sup>

We conduct an additional residence-based validity check in New York City, in which public records provide summary data on where NYPD officers live at the zip code level. We compare the NYPD’s official record of the number of officers who live in a zip code with our smartphone-based estimate of the number of officers that “live” in that same zip code. There is a strong and positive correlation ( $\rho = 0.71$ ) between official NYPD records and our smartphone-based measures. <sup>7</sup>

There is a well-established positive correlation between the fraction of a city population that is Black and the number of sworn police officers per capita (Carmichael and Kent 2014; Stults and Baumer 2007). A basic test of construct validity is whether we observe a similar pattern in our data. Figure 1 plots per capita patrol officers (i.e. smartphones that have at least one “shift”) against the share of Black population in the 21 cities, replicating the positive correlation between the fraction of city residents who are Black and our measure of total officers per capita. Our GPS-based measure of police presence also has significant predictive power on downstream measures of police actions, such as stops and arrests. After adjusting for nonlinearity, the correlation between our measure of police presence and the number of arrests—which we observe in six cities—ranges from 0.46 (Washington) to 0.68 (Austin). Similar positive and significant correlations for police stops for nine cities with publicly available geocoded records are observed as well. <sup>8</sup>

---

<sup>6</sup>Appendix Figure A.3 plots the raw data. The p-value testing whether the slope between the smartphone GPS measures and LEMAS measures of officer racial composition is equal to 1 is 0.529 for Black, 0.611 for White, and 0.098 for Hispanic. The slopes between the two estimates for the share of Asian is significantly different from 1, though Asians account for only 2.5% of the police force across the cities in LEMAS. Table A.1 in the Appendix further reports the correlation conditional on each city’s racial composition.

<sup>7</sup>Figure A.4 in the Appendix plots the zip code level data.

<sup>8</sup>Appendix Figures A.5 and A.6 plot these city by city graphs.



## 3 Results

### 3.1 Neighborhood Correlates of Police Presence

Understanding how police provide services to people from different racial groups is important from both an equity and an efficiency perspective, and our data are uniquely suited to provide new evidence on this issue. Within each neighborhood, we use 2013-2017 American Community Survey (ACS) data to estimate the percent of neighborhood residents who report being in a particular racial or ethnic category. Table A.2 in the Appendix shows summary statistics for police presence measures as well as neighborhood correlates.

Table 1 presents our estimates of the spatial determinants of policing in America’s largest cities. Our smartphone GPS data reveal a strong relationship between the racial and ethnic composition of a neighborhood and police presence. In the largest cities in America, police spend 2.7% more time in places where the fraction of residents who are Black is 10 percentage points higher, 5% more time in places where the share of Hispanic residents is 10 percentage points higher, and 3.9% more time in a place where the share of Asian residents is 10 percentage points higher.<sup>9</sup>

Why do these disparities exist? Differences in where police spend their time can reflect decisions made by individual officers - who ultimately decide where they will go on the job - and department-level directives on patrol assignments. Both involve an assessment, by department or officer, of the residential “need” for police presence in an area. Applicable departmental policies, officer decisions, and residential demand for police presence can all be related to the racial composition of a neighborhood. We use a multivariate OLS regression framework to provide insight into why police may tend to spend more time in places with relatively more Asian, Black, and Hispanic residents.

In column 2, we include city fixed effects. Conditioning on geography differences out any

---

<sup>9</sup> $\text{arsinh}(y) = \ln(y + \sqrt{y^2 + 1}) \approx \ln(2y) = \ln 2 + \ln y$ ; hence the interpretation of  $\beta$  is similar to a log-transformation. 10 pp increase in % Black (Hispanic, Asian) is associated with a  $e^{0.271 \times 0.1} - 1 = 2.7\%$  ( $e^{0.488 \times 0.1} - 1 = 5\%$ ,  $e^{0.379 \times 0.1} - 1 = 3.9\%$ ) increase in police hours.

preference of officials in cities with different residential racial compositions for a particular type of policing, that may contribute to observed disparities (e.g., departments in cities with larger Black populations encouraging officers to make aggressive Terry stops or use predictive policing, see [Meares 2015](#) or [Brayne 2020](#)). City fixed effects also address concerns that our results are driven by a correlation between a city’s racial distribution and the accuracy of our smartphone data. Focusing on variation within cities almost doubles the estimated extra time officers spend in more Black neighborhoods, and reduces the differential policing of more Asian and Hispanic neighborhoods by 26-27%. <sup>10</sup>

We next introduce proxies for residential demand. If officers spend more time in places where there are more people, variation in population density that is correlated with race may contribute to spatial differences in policing. Residents may request that officers respond to crimes, and in particularly disadvantaged neighborhoods, police officers may be one of the few remaining providers of any social service that people need ([Lum 2021](#)). Racial disparities in police presence may therefore stem from racial inequity in the quality of non-policing institutions.

We draw on existing social science literature to approximate components of residential demand for police presence. A lack of educational opportunity and well-paying jobs are established root causes of crime ([Messner and Rosenfeld 1997](#)). Of course, neighborhoods where residents have low incomes but high social capital (i.e. high degrees of social cohesion and community engagement) are places where police rarely need to respond to acts of violence or property destruction ([Sampson and Raudenbush 1999](#)). Following [Martin and Newman \(2015\)](#), we measure social capital using the fraction of 2010 census forms returned by residents. Finally, police officers go where violent crime exists. We estimate the crime-driven demand for policing based on the location of homicides known to the police. While imperfect and sparse, police records of homicides are generally thought to be the most accurate, in the sense that reporting of homicides is unlikely to be as influenced by police presence as

---

<sup>10</sup>In the Online Appendix [D](#), we show that the relationship between exposure to police presence and the composition of the block group that is Black or Hispanic is highest during the middle of an officer’s shift.

reporting of other types of crime, and victimization data suggests that variation in homicides is highly correlated with variation in other crimes (Levitt 1998). We calculate the distance from the neighborhood center to the closest homicide in 2016, treating these rare events as an extreme expression of underlying social issues, implicitly assuming both that crime is spatially clustered and that this distance is negatively correlated with exposure to other types of crime. Additionally, we control for the number of homicides in 2016, by neighborhood, to account for potential variation in higher rates of crime.<sup>11</sup>

In column 3 of Table 1, we condition our estimates of local police presence in different types of U.S. neighborhoods on measures of density, socioeconomics, social cohesion, and violence. Differential residential demand for police presence, some of which is created by decisions made in other policy domains, explains approximately 33% of the disparate exposure of people living in relatively Black neighborhoods, 31% of the disparate exposure of people living in relatively Hispanic neighborhoods, and can explain all of the additional exposure of people living in relatively Asian neighborhoods—even suggesting that more Asian neighborhoods have less police presence than one might expect based on social conditions.<sup>12</sup> The residual correlation between racial composition and police presence in column 3 reflects decisions at the police command, and officer level.

Diversifying the officer ranks is one city-level policy that is central to many police reform efforts. With this in mind, we compare how disparities in police presence vary with the racial composition of a city’s police force. We do this in two ways: including the mean-centered interaction between the share of Black residents and the share of police officers that are Black in column 4, and interactions with both the share of police supervisors and patrol officers that are Black in column 5.<sup>13</sup> Column 4 suggests the additional exposure to police in Black

---

<sup>11</sup>Alternate measures of demand for policing, specifically using additional years of homicide data and 311 calls for service in New York City (Shah and LaForest 2021) are explored in the Online Appendix B - none lead to substantively different conclusions.

<sup>12</sup>In the Online Appendix C, we show that our findings are qualitatively identical when we model police presence during non-working hours (excluding weekday 9 am - 5 pm), and in New York City when we exclude census block groups in tourist destinations. In both of these situations, the demographics of residents may differ from the demographics of the ambient population.

<sup>13</sup>Appendix Figure A.8 reveals substantial cross city variation in the share of Black police officers and

neighborhoods is larger in cities with a larger share of Black officers; while this cross-city comparison is not necessarily inconsistent with existing work, it stands in contrast to single city studies finding that Black officers spend less time in Black neighborhoods (Ba et al., 2021). Further, column 5 implies that, conditional on the composition of patrol officers, there is less police presence in Black neighborhoods when more front-line supervisors are Black.

More specifically, relative to a city with the mean number of Black officers and supervisors and conditioning on social conditions, a 10 percentage point increase in the number of Black officers would mean that a 1 percentage point increase in the share of residents who are Black is associated with a 0.53% increase in police presence.<sup>14</sup> If there were a simultaneous 37.6 percentage point increase in Black supervisors, we would observe no relationship between the fraction of neighborhood residents who are Black and the time police spend in that neighborhood.<sup>15</sup> While correlational in nature, our findings suggest that efforts to hire more Black police officers, without parallel efforts to retain and promote those officers, may not reduce disparities in how the public is policed.

### 3.2 Cross-city variation in correlates of police presence

Our findings suggest substantial differences in the level of ambient police presence in non-White neighborhoods across the United States, and this difference is largest in Black (relative to White) neighborhoods. Given this, and the long and fraught history of the policing of Black people in the United States, in this section we focus on police presence in relatively Black versus relatively White neighborhoods. First, in Figure 2, we show police presence in neighborhoods with the greatest share of Black and White residents, respectively, to highlight the range of disparities in Black-White neighborhood policing across major US cities. There

---

supervisors, and a meaningful difference between the share of Black officers and supervisors, despite a high correlation between the two measures.

<sup>14</sup>  $e^{(0.01 \times (0.307 + 2.207 \times 0.1))} - 1 = 0.53\%$

<sup>15</sup>  $.307 + .2207 - (1.403 \times .376) = 0$

is little difference in the amount of time that police spend in the “most White” and “most Black” neighborhoods in Boston, but over 100 more hours of total policing in the “most Black” neighborhoods in Charlotte than in the “most White.”

Of course, these disparities can have many sources. In Figure 3, we plot, for each city, how much of the spatial variation in police presence can be explained by spatial variation in our proxies for “demand” for police, and how much the explanatory power of our models increases when we add controls for racial composition. This shows the extent to which Black-White disparities in exposure to police can persist even when considering spatial differences in socioeconomic status—which may reflect historical and contemporary race-based social and economic inequality. We document substantial variation across cities in the role of this structural inequality in explaining policing disparities. For example, while Figure 2 reveals large differences in the ambient police exposure of Charlotte residents in the most Black and most White neighborhoods, Figure 3 reveals that spatial disparities in socioeconomic status explain almost all of these differences. These structural disparities in Charlotte are city-level issues that cannot be addressed solely by the city’s police department. In contrast, racial disparities in police presence are absolutely smaller in Austin, but incorporating Black, Hispanic, and Asian residential patterns increases the amount of spatial variation in police presence that we can explain in that city by 27%. This suggests substantially more scope for changes in police policy to reduce criminal justice disparities in Austin, TX.<sup>16</sup>

### 3.3 Using Police Presence to Understand Police Enforcement

The empirical observation that police are more ambiently present in Non-White neighborhoods provides support for the construct validity of our data, as this correlation has been repeatedly demonstrated at the city level (Carmichael and Kent 2014). When taken in the context of existing qualitative and legal scholarship on modern policing, this also raises equity concerns.

---

<sup>16</sup>Appendix Figure A.9 also plots the city-specific estimate of Black-White disparity.

To quantify the extent to which racial disparities in upstream police presence are associated with disparities in one consequential downstream law enforcement action—arrest—we create three neighborhood-level measures: how much time officers spend in a given neighborhood, how many arrests are made in that neighborhood, and how many arrests are made per hour of police presence. Our measure of police presence can therefore distinguish between two very different sources of racial disparities in arrests: variation in ambient police presence that is correlated with race, and differences in behavior of officers across different neighborhood contexts.

Consistent with studies of downstream measures of policing, Table 2 confirms that in six cities for which we have both police presence and arrest data (New York City, Los Angeles, Chicago, Dallas, Austin, Washington), officers spend more time, and make more arrests in neighborhoods with more Black residents than the typical neighborhood in each city. Column 1 shows that the Black-White disparity in police presence is 20% smaller, the Hispanic-White disparity 55% smaller, and the Asian-White disparity 20% larger, in this set of cities that choose to make geocoded arrest data public.

While column 2 shows that officers make approximately 21% more arrests in neighborhoods where the share of residents who are Black is 10 percentage point higher, in column 3 we show that they make almost 13% more arrests per hour present.<sup>17</sup> We find that our proxies for neighborhood demand for police do explain part of the increased number of arrests in more Black neighborhoods, but in this sub-sample, they do not explain the increased police presence—in fact, the residual disparities increase. Whatever the source, this disparity in the propensity of an officer to make an arrest in more Black neighborhoods, while keeping other socioeconomic variables constant, explains less than half of the residual neighborhood disparity in the total number of arrests made.<sup>18</sup> This implies that the added time that police

---

<sup>17</sup>The semi-elasticity is approximately equal to  $e^{1.909 \times 0.1} - 1 = 21\%$  ( $e^{1.231 \times 0.1} = 13\%$ ) for a 10 percentage point increase in % Black (% Hispanic).

<sup>18</sup>Specifically, the elasticity of arsinh-linear model is  $\beta \bar{x} \sqrt{\frac{y^2 + 1}{y^2}} \approx \beta \bar{x}$ . Differences in the propensity of officers to make an arrest while in a relatively more Black neighborhood explain 44.6% ( $\frac{\beta_2 \bar{x}}{\beta_1 \bar{x}} = \frac{0.619}{1.387}$ ) of the neighborhood arrest disparity, and the remaining 55.4% is explained by differences in the police

spend in Black neighborhoods may be a central source of Black-White disparities in arrests, in addition to an officer’s decision in a particular encounter. It is outside the scope of this paper to evaluate the welfare implications of this empirical fact, which could be due to over- (or under-)policing, police using different standards to determine if people in different groups are suspicious enough to warrant an arrest, differences in unobserved criminal activity, or to differences in how police officers spend time in these neighborhoods.<sup>19</sup> Consistent with Meares (2015), our results suggest that in order to reduce disparities in criminal justice, reducing the scope for racial bias both in officers’ decisions during civilian encounters and in departmental directives detailing where officers go and who they surveil may be warranted.

## 4 Conclusion

We conclude by noting that a positive correlation in the provision of policing and the concentration of Black residents stands in contrast with documented spatial patterns of other institutional investment in neighborhoods with concentrated Black populations, which Derenoncourt (2022) also documents at the city level. Census tracts where more of the residential population is Black are less, not more, likely to have a large grocery store, nearby hospital, or local banking services (Walker et al. 2010; DeYoung et al. 2008; Lieberman-Cribbin et al. 2020; Yearby 2018). During the 2016 election, Chen et al. (2019) found that voting lines moved more slowly in places with larger Black populations, suggesting under-investment in polling services in places where we observe larger investments in ambient policing.

Our data are well suited to further research on policing in the United States. First, 

---

presence across more Black versus more White neighborhood. Conditioning on socioeconomic characteristics, additional police presence in relatively more Hispanic neighborhoods explain 59.6% ( $1 - \frac{0.466}{1.154}$ ) of the Hispanic-White neighborhood arrest disparity. Online Appendix Table A.7 also reveals a highly similar pattern regarding stop disparities.

<sup>19</sup>Police using different decision rules in more and less Black neighborhoods, while an axiomatic example of discrimination, is not necessarily illegal; Illinois v Wardlow, 528 U. S. 119 (2000) established that officers can use the predetermined designation of an area as “high crime” in determining how likely it is that someone has (or is) engaged in crime, creating a legal basis for a stop. If places with more Black or Hispanic residents are more likely to be known to police as “high crime” places, then this would lower the standard of individualized suspicion needed to make a constitutionally permissible stop.

smartphone location data provide insight into officer presence in communities that traditional measures of policing cannot fully capture. Measuring officer presence informs estimates of which communities are at risk of more serious police encounters, like arrest or the use of lethal force. Second, our smartphone location data do not depend on software purchased by or developed for a particular policing agency, allowing us to map officer locations in cities across the United States using a consistent methodology. This is an advantage over technologies like Automated Vehicle Locators and body cameras, because it provides enhanced visibility into the unreported and highly discretionary activities of police officers at work. Finally, data on where officers spend their patrol time grants researchers and practitioners new abilities to understand patterns in police presence and track the implementation of departmental policies that shape the provision of public safety.



## References

- Ba, Bocar A., Dean Knox, Jonathan Mummolo, and Roman Rivera (2021) “The role of officer race and gender in police-civilian interactions in Chicago,” *Science*, 371 (6530), 696–702.
- Banaji, Mahzarin R, Susan T Fiske, and Douglas S Massey (2021) “Systemic racism: individuals and interactions, institutions and society,” *Cognitive research: principles and implications*, 6 (1), 1–21.
- Banks, R Richard, Jennifer L Eberhardt, and Lee Ross (2006) “Discrimination and implicit bias in a racially unequal society,” *California Law Review*, 94 (4), 1169–1190.
- Brayne, Sarah (2020) *Predict and surveil: Data, discretion, and the future of policing*: Oxford University Press, USA.
- Carmichael, Jason and Stephanie Kent (2014) “The persistent significance of racial and economic inequality on the size of municipal police forces in the United States, 1980–2010,” *Social Problems*, 61, 259–282.
- Chen, M Keith, Judith A Chevalier, and Elisa F Long (2021) “Nursing home staff networks and COVID-19,” *Proceedings of the National Academy of Sciences*, 118 (1), e2015455118.
- Chen, M Keith, Kareem Haggag, Devin G Pope, and Ryne Rohla (2019) “Racial disparities in voting wait times: evidence from smartphone data,” *Review of Economics and Statistics*, 1–27.
- Derenoncourt, Ellora (2022) “Can You Move to Opportunity? Evidence from the Great Migration,” *American Economic Review*, 112 (2), 369–408, [10.1257/aer.20200002](https://doi.org/10.1257/aer.20200002).
- DeYoung, Robert, W Scott Frame, Dennis Glennon, Daniel P McMillen, and Peter Nigro (2008) “Commercial lending distance and historically underserved areas,” *Journal of Economics and Business*, 60 (1-2), 149–164.

- Geller, Amanda, Jeffrey Fagan, Tom Tyler, and Bruce G. Link (2014) “Aggressive Policing and the Mental Health of Young Urban Men,” *American Journal of Public Health*, 104 (12), 2321–2327.
- Goel, Sharad, Maya Perelman, Ravi Shroff, and David Sklansky (2017) “Combatting Police Discrimination in the Age of Big Data,” *New Criminal Law Review: An International and Interdisciplinary Journal*, 20, 181–232, [10.1525/nclr.2017.20.2.181](https://doi.org/10.1525/nclr.2017.20.2.181).
- Hinton, Elizabeth (2016) *From the War on Poverty to the War on Crime: The Making of Mass Incarceration in America*: Harvard University Press.
- Hoekstra, Mark and CarlyWill Sloan (2022) “Does Race Matter for Police Use of Force? Evidence from 911 Calls,” *forthcoming, American Economic Review*.
- Levitt, Steven D (1998) “The relationship between crime reporting and police: Implications for the use of Uniform Crime Reports,” *Journal of Quantitative Criminology*, 14 (1), 61–81.
- Lieberman-Cribbin, Wil, Stephanie Tuminello, Raja M Flores, and Emanuela Taioli (2020) “Disparities in COVID-19 testing and positivity in New York City,” *American Journal of Preventive Medicine*, 59 (3), 326–332.
- Lum, Cynthia (2021) “Perspectives on Policing,” *Annual Review of Criminology*, 4, 19–25.
- Martin, David C and Benjamin J Newman (2015) “Measuring aggregate social capital using census response rates,” *American Politics Research*, 43 (4), 625–642.
- Mastrobuoni, Giovanni (2019) “Police disruption and performance: Evidence from recurrent redeployments within a city,” *Journal of Public Economics*, 176, 18–31.
- Meares, Tracey L (2015) “Programming errors: Understanding the constitutionality of stop-and-frisk as a program, not an incident,” *University of Chicago Law Review*, 82, 159.

- Messner, Steven F and Richard Rosenfeld (1997) “Political restraint of the market and levels of criminal homicide: A cross-national application of institutional-anomie theory,” *Social Forces*, 75 (4), 1393–1416.
- OJJDP Statistical Briefing Book (2019) *Estimated number of arrests by offense and race*, [https://www.ojjdp.gov/ojstatbb/crime/ucr.asp?table\\_in=2](https://www.ojjdp.gov/ojstatbb/crime/ucr.asp?table_in=2).
- Pierson, Emma, Camelia Simoiu, Jan Overgoor et al. (2020) “A large-scale analysis of racial disparities in police stops across the United States,” *Nature Human Behaviour*, 4 (7), 736–745.
- Rios, Victor (2011) *Punished: Policing the Lives of Black and Latino Boys*: New York University Press.
- Rucker, Julian M and Jennifer A Richeson (2021) “Toward an understanding of structural racism: Implications for criminal justice,” *Science*, 374 (6565), 286–290.
- Sampson, Robert J and Stephen W Raudenbush (1999) “Systematic social observation of public spaces: A new look at disorder in urban neighborhoods,” *American Journal of Sociology*, 105 (3), 603–651.
- Shah, Anuj K. and Michael LaForest (2021) “Knowledge About Others Reduces One’s Own Sense of Anonymity,” *Working Paper*.
- Sharkey, Patrick (2018) *Uneasy Peace: The Great Crime Decline, the Renewal of City Life, and the Next War on Violence*: W. W. Norton and Company.
- Stults, Brian J and Eric P Baumer (2007) “Racial context and police force size: Evaluating the empirical validity of the minority threat perspective,” *American Journal of Sociology*, 113 (2), 507–546.
- Vomfell, Lara and Neil Stewart (2021) “Officer bias, over-patrolling and ethnic disparities in stop and search,” *Nature Human Behaviour*, 5 (5), 566–575.

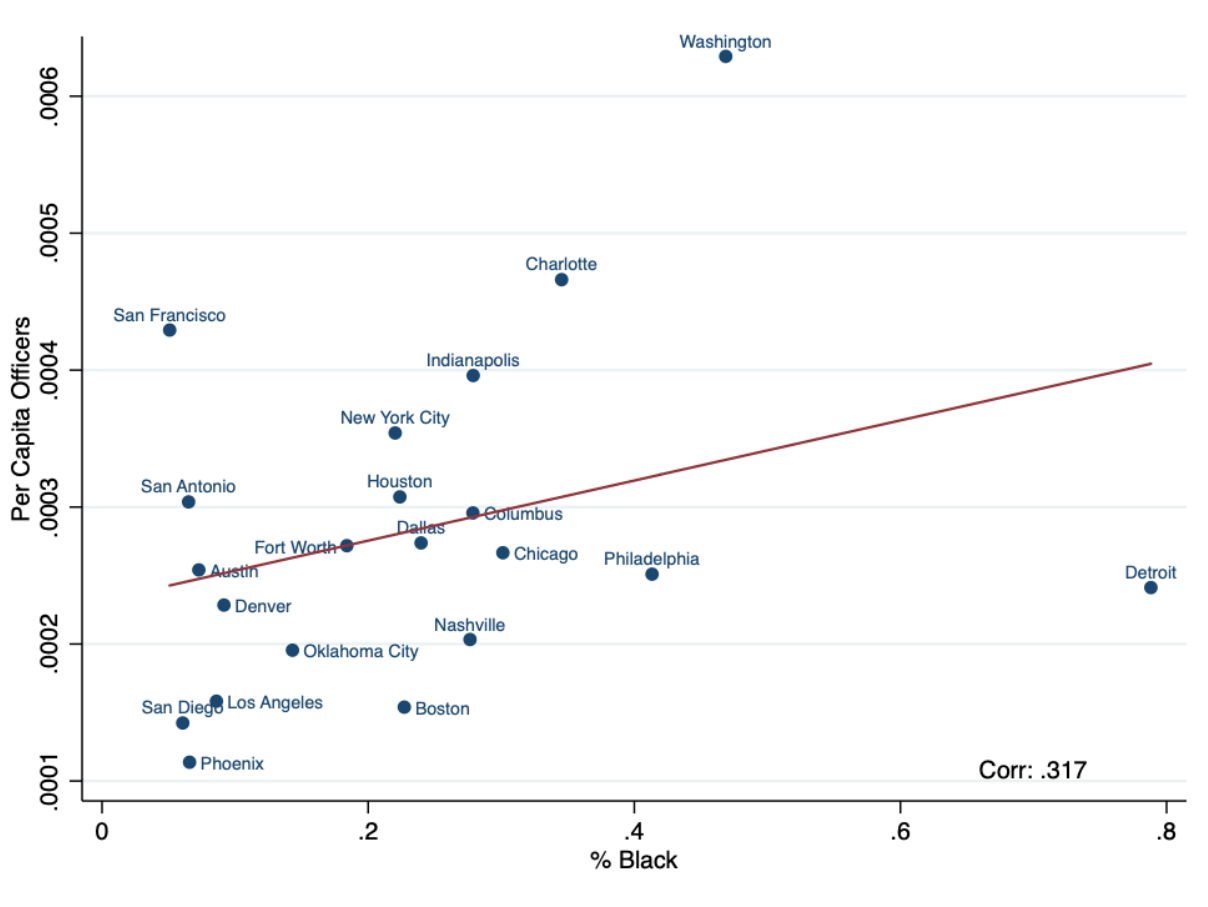
Walker, Renee E, Christopher R Keane, and Jessica G Burke (2010) “Disparities and access to healthy food in the United States: A review of food deserts literature,” *Health & Place*, 16 (5), 876–884.

Weisburd, Sarit (2021) “Police presence, rapid response rates, and crime prevention,” *Review of Economics and Statistics*, 103 (2), 280–293.

Yearby, Ruqaiijah (2018) “Racial disparities in health status and access to healthcare: the continuation of inequality in the United States due to structural racism,” *American Journal of Economics and Sociology*, 77 (3-4), 1113–1152.

# Figures and Tables

Figure 1: Correlation Between % Black and Officers per capita in a City



*Notes:* Per capita officers is defined as the number of likely patrol officers on “shift” (identified with smartphone data) divided by the city population (2013-2017 American Community Survey estimate). We identify patrol officers on “shift” by looking for a specific pattern in smartphones that visit police stations at least 5 days in a month: Leaving “home”, traveling to a police station, moving around the city (without returning home), returning to the police station, and then going home. The correlation coefficient between the two measures is reported.

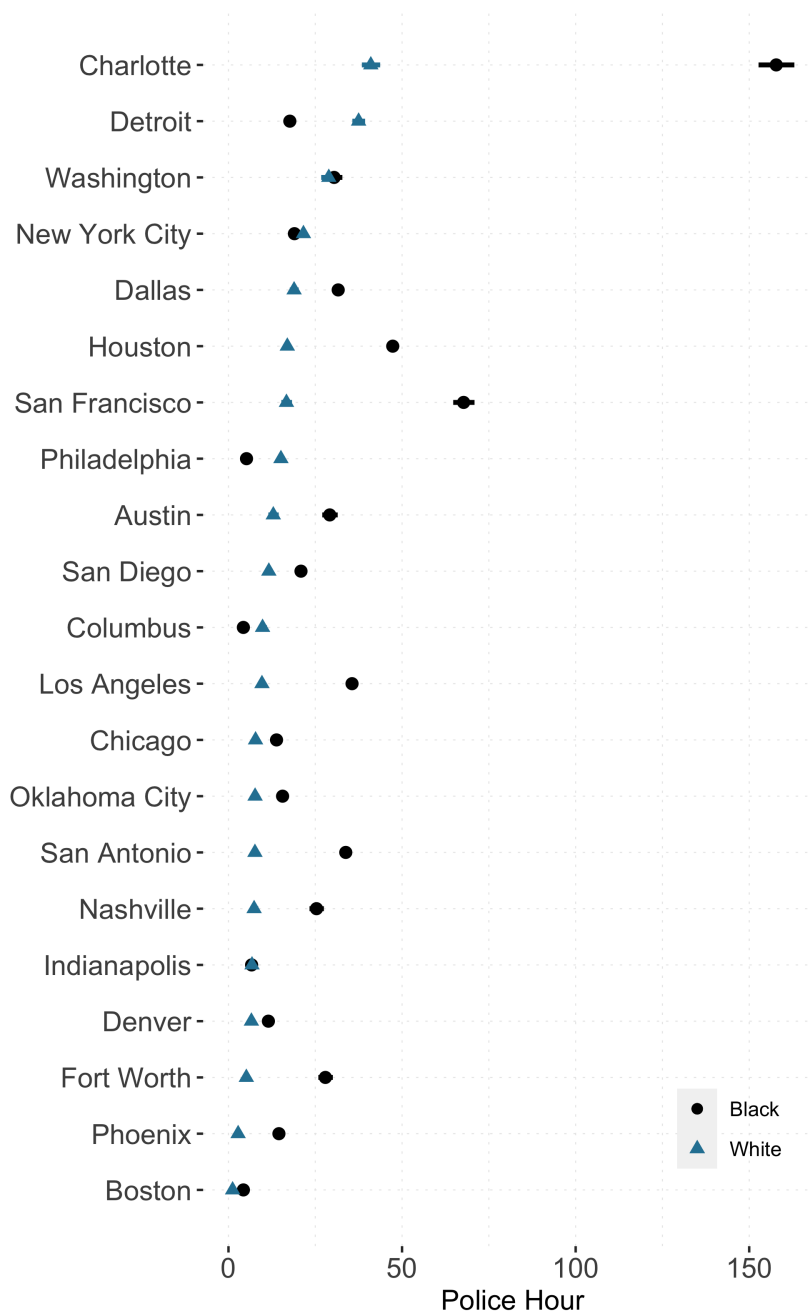


Figure 2: Police Exposure in Blackest and Whitest Neighborhoods

Notes: This figure plots the average police hours observed in the Blackest (Whitest) neighborhoods in a city, defined as the block groups where share of Black (White) residents is over the 95th percentile of the city’s distribution. The cities are ordered by police presence in the whitest neighborhoods.

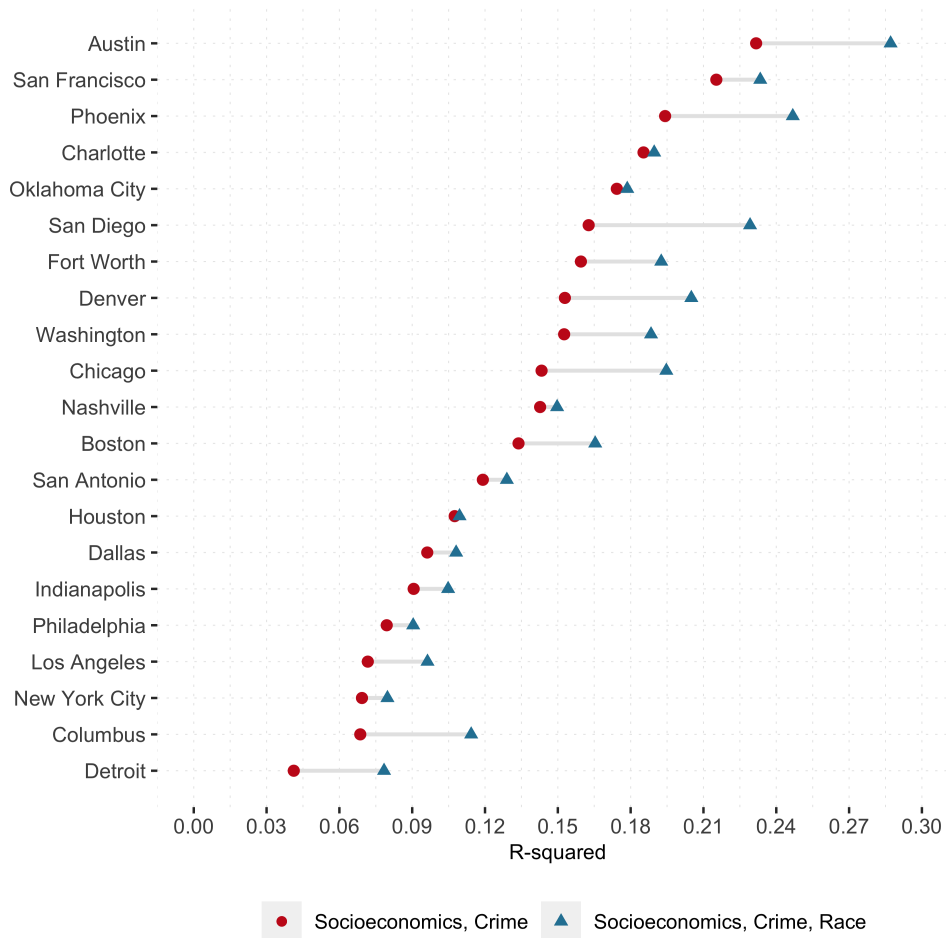


Figure 3: Variance of Police Hours Explained by Socioeconomics, Crime, and Race

Notes: This figure reports the R-squares of the following two OLS regressions for each city: (1)

$$\text{arsinh}(\text{Hour}_i) = \beta_0 + \beta_1 \text{Socioeconomics}_i + \beta_2 \text{Crime}_i + \epsilon_i, \text{ and (2)}$$

$\text{arsinh}(\text{Hour}_i) = \beta_0 + \beta_1 \text{Socioeconomics}_i + \beta_2 \text{Crime}_i + \beta_3 \text{Race}_i + \epsilon_i$ . *Socioeconomics* include log population, % college graduates, median household income, census form return rate. *Crime* include distance to nearest homicide and homicide count in 2016. *Race* include percent Black, percent Hispanic and percent Asian in the block group.

Table 1: Disparities in Neighborhood Police Exposure

VARIABLES	(1)	(2)	(3)	(4)	(5)
	Police Exposure in a Census Block Group: $\text{arsinh}(\text{Hours})$				
% Black	0.271*** (0.0332)	0.493*** (0.0353)	0.329*** (0.0482)	0.320*** (0.0510)	0.307*** (0.0531)
BG % Black X Police: % Black				0.715* (0.298)	2.207* (0.901)
BG % Black X Supervisor: % Black					-1.403+ (0.817)
% Hispanic	0.488*** (0.0348)	0.362*** (0.0367)	0.248*** (0.0569)	0.224*** (0.0597)	0.203*** (0.0607)
% Asian	0.379*** (0.0735)	0.276*** (0.0783)	-0.0650 (0.0827)	-0.0941 (0.0843)	-0.0874 (0.0846)
Log Population			0.414*** (0.0212)	0.428*** (0.0220)	0.454*** (0.0226)
% College Graduates			1.105*** (0.0683)	1.162*** (0.0708)	1.187*** (0.0716)
Median Household Income (1K)			-0.00431*** (0.000397)	-0.00425*** (0.000406)	-0.00406*** (0.000409)
Census Form Return Rate			-1.275*** (0.127)	-1.341*** (0.133)	-1.402*** (0.136)
Distance to nearest 2016 homicide (km)			-0.118*** (0.00668)	-0.113*** (0.00738)	-0.109*** (0.00760)
Homicide Count 2016			0.202*** (0.0201)	0.203*** (0.0209)	0.200*** (0.0214)
Constant	2.308*** (0.0168)	2.353*** (0.0168)	0.411* (0.179)	0.372* (0.186)	0.194 (0.190)
Observations	23,682	23,682	22,521	20,961	20,112
R-squared	0.008	0.106	0.167	0.154	0.159
City FE	No	Yes	Yes	Yes	Yes

*Notes:* This table presents OLS estimates of exposure disparities among census block groups  $i$  (BGs) across (Column 1, 2, 4, 5) and within 21 of the largest US cities (Column 3, 6):  $\text{arsinh}(\text{Hour}_i) = \beta_0 + \beta_1 X_i + \epsilon_i$ .

Coefficient estimates for all variables in  $X_i$  are reported in the table. The dependent variable is police hours observed in a BG (excluding pings moving faster than 50 mph), transformed into arsinh values. All race variables (including neighborhood racial composition, Police: % Black and Supervisor: % Black) are mean-centered. Household income is measured in thousands of dollars, census return rates range from 0-1. Robust standard errors are reported in parentheses. Results are qualitatively and quantitatively similar to running all regressions with log dependent variable and dropping zero-valued observations, or clustering at the tract level, and are available on request. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ , +  $p < 0.1$



Table 2: Disparities in Neighborhood Police Exposure and Downstream Disparities

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	arsinh(Hours)	arsinh(Arrests)	arsinh(Arrests/Hour)	arsinh(Hours)	arsinh(Arrests)	arsinh(Arrests/Hour)
% Black	0.398*** (0.0459)	1.909*** (0.0398)	1.231*** (0.0404)	0.638*** (0.0642)	1.387*** (0.0581)	0.619*** (0.0584)
% Hispanic	0.160*** (0.0467)	1.611*** (0.0421)	1.118*** (0.0397)	0.519*** (0.0750)	1.154*** (0.0683)	0.466*** (0.0660)
% Asian	0.329*** (0.0933)	0.712*** (0.0851)	0.210** (0.0714)	0.355*** (0.100)	0.219* (0.0913)	-0.190* (0.0814)
Log Population				0.515*** (0.0311)	0.500*** (0.0267)	-0.0618** (0.0223)
% College Graduates				1.499*** (0.0921)	0.692*** (0.0850)	-0.732*** (0.0768)
Median Household Income (1K)				-0.00277*** (0.000503)	-0.00401*** (0.000464)	-0.000736+ (0.000390)
Census Form Return Rate				-0.661*** (0.169)	-1.799*** (0.145)	-0.860*** (0.144)
Distance to nearest 2016 homicide (km)				-0.137*** (0.0134)	-0.157*** (0.0123)	0.00114 (0.0116)
Homicide Count 2016				0.230*** (0.0279)	0.355*** (0.0224)	0.0991*** (0.0228)
Observations	12,748	12,748	12,705	12,098	12,098	12,059
R-squared	0.049	0.239	0.197	0.123	0.326	0.213
City FE	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:* This table presents OLS estimates of disparities in exposure, arrests, and arrests per hour among census block groups  $i$  (BGs) across 6 cities:  $arsinh(Y_i) = \beta_0 + \beta_1 X_i + \epsilon_i$ . Coefficient estimates of all variables in  $X_i$  are reported in the table. The six cities with publicly available geocoded arrest data are: New York City, Los Angeles, Chicago, Dallas, Austin, Washington. The dependent variables ( $Y_i$ ) are: police hours observed in a BGs (excluding pings moving faster than 50 mph, mean 26.9), number of arrests in that BG (mean 40.1), and the ratio of those two measures (mean 10.2), all transformed into arsinh values. Household income is measured in thousands of dollars, census return rates range from 0-1. Robust standard errors are reported in parentheses. Results are qualitatively and quantitatively similar to running all regressions with log dependent variables and dropping zero-valued observations, or clustering at the tract level, and are available on request. \*\*\* p<0.001, \*\* p<0.01, \* p<0.05, + p<0.1