

Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial

P. Jeffrey Brantingham*

and

Matthew Valasik[†]

and

George O. Mohler[‡]

June 3, 2018

Abstract

Racial bias in predictive policing algorithms has been the focus of a number of recent news articles, statements of concern by several national organizations (for example the ACLU and NAACP), and simulation based research. There is reasonable concern that predictive algorithms encourage directed police patrols to target minority communities with discriminatory consequences for minority individuals. However, to date there have been no empirical studies on the bias of predictive algorithms used for police patrol. Here we test for such biases using arrest data from the Los Angeles predictive policing experiments. We find that there were no significant differences in the proportion of arrests by racial-ethnic group between control and treatment conditions. We find that the total numbers of arrests at the division level declined or remained unchanged during predictive policing deployments. Arrests were numerically higher at the algorithmically predicted locations. When adjusted for the higher overall crime rate at algorithmically predicted locations, however, arrests were lower or unchanged.

Keywords: crime forecasting; civil liberties; police bias

*UCLA Department of Anthropology, Los Angeles, CA 90095, USA.

[†]Louisiana State University Department of Sociology, Baton Rouge, LA 70803, USA.

[‡]Indiana University-Purdue University Indianapolis Department of Computer and Information Science, Indianapolis IN 46202, USA.

1 Introduction

Place-based predictive policing is based on two core ideas: (1) mathematical forecasting methods can be used to anticipate future crime risk in narrowly prescribed geographic areas; and (2) the delivery of police resources to those prediction locations disrupts the opportunity for crime [42, 10]. Randomized controlled experiments of predictive policing conducted in Los Angeles provided evidence that algorithmic methods not only predict two-times as much crime as existing best practice, but also double the amount of crime prevented [41]. While this treatment effect can be measured in the field, the specific mechanisms by which predictive policing delivers greater crime reduction are not immediately obvious.

The prevailing view, derived from experiments in hot spot policing [55, 11], is that the presence of police in a given place removes opportunities for crime even without any direct contact with potential offenders [56, 62, 36]. This general deterrent effect persists for some time after police have departed [34, 56], and appears to diffuse into nearby areas where the police were not concentrating their efforts [14, 64, 60]. General deterrence is not the only mechanism by which crime might be prevented by police patrol, however. Direct interference via stops, searches, detentions short of arrest, and arrest, may prevent crime by physically incapacitating potential offenders [54, 63]. This use of selective incapacitation may have immediate effects on crime [65], especially if prolific offenders are the ones being arrested. Incapacitation may have longer term effects if those prolific offenders are subsequently removed from the community.

Considerable evidence suggests that explicit and implicit bias can have a major impact on *who* gets stopped, searched and detained. Reasonable concern therefore exists that predictive policing can exacerbate such biases and reinforce any tendency for police to target minority individuals and communities [23]. Such concern exists even if the forecasting methods used to drive predictive policing refrain from incorporating data that would be an explicit source of bias. If predictive policing indirectly exacerbates bias, any crime control benefits would need to be weighed in terms of their discriminatory costs. In the worst case, documented benefits might be derived solely from bias induced by predictions. In other words, predictions absent such bias would yield no crime control benefits at all. Here we seek to evaluate whether predictive policing leads to patterns of arrest biased against minority individuals.

2 Predictive Policing and Racial Bias

Research has demonstrated that a racial bias exists in the business of policing including the racial profiling of vehicles [5, 17, 20, 28, 30, 40, 46, 48, 58, 61],

pedestrian stops [18, 19, 25, 26, 27, 49, 59], traffic tickets [16], drug enforcement and arrests [7, 9, 33, 38, 57], use of force [45, 13, 35, 51], and even in the decision to shoot white or black criminal suspects while in a training simulator [24, 47]. While the mechanisms driving these observed patterns of racial disparity (i.e., racial profiling, stereotyping/cognitive bias, deployment, racial animus/prejudice) remain difficult to disentangle, as [61] attests, there is little doubt that racial disparities in policing outcomes do exist.

Racial bias of predictive policing algorithms has been the focus of a number of recent news articles [3, 2], and concerns have been raised by several national organizations (for example the ACLU and NAACP) [1] and recent research articles [12, 23, 31]. In regards to place-based predictive policing methods that forecast a time and location where a crime may occur, the concern is that racially biased police practices may be directed towards some areas rather than others [23]. In addition, knowing that they are in a prediction area may heighten the awareness of police officers in ways that amplify biases [21]. That is, a minority individual observed in a prediction area may be more likely to be subject to biased police actions than the same individual observed outside of a prediction area.

Lum and Isaac [37] conduct a simulation study of predictive policing focused on drug arrests in Oakland, CA. Their goal is to ascertain if racially biased outcomes are possible, or even amplified, with place-based predictive policing methods. The algorithm Lum and Isaac analyze is a space-time Hawkes process that was the method used in the Los Angeles Predictive Policing Experiment [41]. In particular, given a city grid indexed by n , the probabilistic rate $\lambda_n(t)$ of events in cell n at time t is determined by,

$$\lambda_n(t) = \mu_n + \sum_{t_n^i < t} \theta \omega e^{-\omega(t-t_n^i)}, \quad (1)$$

where t_n^i are the times of events in cell n in the history of the process, μ_n is a baseline rate of events, and $\theta \omega e^{-\omega t}$ reflects the increase in risk following a recent crime. Lum and Isaac show that if the events t_n^i correspond to racially biased drug arrests, then $\lambda_n(t)$ will increase after an arrest, leading to more police resources deployed to cell n in the future. Thus a feed back loop may be possible where more arrests then occur in cell n , leading to a further increase in λ_n . A similar concern is raised by Ferguson [22], who notes that arrests in a prediction area ‘memorializes’ that location as ‘hot’, which guarantees that that it will show up again as a prediction area producing further arrests. Selbst [52] also warns that racially biased outcomes may become entrenched in place-based predictive policing given data collected via discriminatory policing practices. Constitutional policing tenets such as reasonable suspicion and probable cause may offer little protection against bias embedded in the data [15, 53]. Though all of these studies deal with

hypothetical scenarios or thought experiments, they succeed in demonstrating that careful attention needs to be paid to whether predictive policing produces biased arrests [12, 15, 31, 43].

In practice, the majority of hotspot and place-based predictive policing algorithms focus not on arrests, but on crimes predominantly reported to the police by the public (e.g., robbery, burglary, assault) [8, 22, 32, 41]. Thus the goal is to send police resources to areas where crimes have been reported by victims, thus preventing future crimes in those areas. While a feedback loop for reported crime may be possible, in this case the self-reinforcement is towards places where citizens are placing calls for service. Therefore we focus on the question of whether predictive policing produces arrests biased against minorities when the inputs to the system are reported crime incidents, rather than arrests. We run a set of hypotheses tests on empirical arrests recorded during the Los Angeles predictive policing experiment. We ask three related questions: (1) Did arrest of minority individuals differ between control and treatment conditions in test divisions? (2) Did arrest rates overall differ between control and treatment conditions in test divisions? and (3) Did the rate of arrests per crime differ across treatment and control conditions.

3 Predictive Policing Experiments in Los Angeles

A randomized controlled trial of predictive policing was conducted in three divisions of the Los Angeles Police Department (LAPD) between November 2011 and January 2013. The three participating LAPD divisions were Foothill (FH), North Hollywood (NH) and Southwest (SW). Only a brief outline of the experiment is presented here. Additional details of the algorithmic procedures, experimental design and main effects are presented in [41].

Each day of the experiment police patrol officers were handed patrol maps with twenty target areas marked as 500 x 500 foot boxes. Officers were informed that the target areas were locations where the risk of crime was highest for their shift. They were encouraged to patrol target areas during any available discretionary time. What officers did not know was that the mission maps distributed to them each day were designed either by an algorithmic forecasting method [see 42, 41], or by an analyst from within the division using all of the technological and intelligence assets at their disposal. Which mission map officers received on any given day was randomized creating a treatment condition (algorithmic forecast) and control condition (analyst forecast). In this repeated-measures experimental design, treatment days were considered exchangeable with control days [41].

The outcome of interest was the difference in reported crime between control

and treatment days. The crime types targeted were burglary, car theft and burglary theft from vehicle (BTFV). Historically, these crime types account for as much as 60% of the serious crime in the City of Los Angeles. In addition to this outcome measure, we collected information on the amount of time police officers spent in prediction areas under each of the experimental conditions [41]. Officers used their in-car computer terminals to register when they were entering and exiting prediction locations. This “dosage” was aggregated by day for a total amount of time (in minutes) spent in prediction locations.

Across the three test divisions, patrol officers using the algorithmic predictions produced an average 7.4% drop in crime at mean patrol dosage. By contrast, use of the best-practice predictions produced an average 3.5% drop in crime at mean patrol dosage. Individually, the decrease in crime associated with algorithmic predictions was statistically significant, while that with best-practice predictions was not. The evidence presented in [41] suggests that that police patrol, when influenced by accurate predictions about the timing and location of crime, may provide some additional crime deterrence value. However, the standard errors of the estimates are relatively large making the slopes difficult to distinguish statistically and the precision on these rates is insufficient to conclude that the algorithmic forecasting crime rate drop is greater than the analyst crime rate drop.

3.1 Data

3.1.1 Defining Experimental Periods

Predictive policing experiments in each of the three LAPD divisions started at different times and ran for different total durations (Figure 1). The experiment in Foothill Division ran for 172 days from November 7, 2011 to April 27, 2012. In North Hollywood, the experiment ran for 167 days from March 31, 2012 to September 14, 2012 and in Southwest Division, the experiment ran for 239 days between May 16, 2012 to January 10, 2013. Total crime and arrests counts differ as a function of experimental duration in each division.

3.1.2 Defining Control and Treatment Days

Control and treatment missions were designed independently, but in parallel each day of the experiment. Recall that treatment missions were based on algorithmic forecasting, while the control missions were based on existing best practice of analysts. Once mission designs were finalized, a control or treatment mission was chosen randomly for deployment. This randomization was done independently each day for each division taking part in the experiment. On occasion, the analyst was not present on a randomly designated control day and therefore control

missions were not available for those days. We exclude treatment days from these days to ensure fair comparison. In Foothill Division, there were a total of 124 test days with successful random assignment, after discarding days on which the analyst was not present to design control missions. The 124 test days were evenly divided with 62 control and 62 treatment days. There were 152 total test days in North Hollywood Division. These included 82 control and 70 treatment days. In Southwest Division, there were 234 total days, including 117 control and 117 treatment days.

3.1.3 Defining Arrests

An arrest is generally understood to mean the taking into custody of an individual by the police given probable cause that a violation of the law has occurred. An arrest, as recorded by the LAPD, should not be conflated with other down-stream processes of the criminal justice system. An arrest does not imply booking, continued detention, nor whether those individuals are ultimately prosecuted for a crime. Arrests also should not be conflated with contacts between the public and police that did not result in arrest, even if such contacts were contentious. In general, police can exercise many alternatives to arrest in seeking to enforce laws and ensure order including behavioral directives, warnings and brief detention without arrest. On average, the LAPD makes about 1.5 million public contacts per year, about 24,000 of these contacts (1.6%) are arrests [6]. Here arrests are taken at face value, without considering anything beyond the official record that an individual was taken into custody.

We do not distinguish between arrests for different types of crimes. In 2012, the LAPD made arrests under 520 different criminal codes representing 25 broad classes of crimes such as aggravated assault, robbery, burglary and larceny. Our primary focus is on whether the practice of policing introduces new biases into arrest patterns, not whether bias might be differentially present in arrests for different types of crimes.

3.1.4 Defining Racial-Ethnic Groups

The LAPD collects demographic information as part of the arrest process including age, sex and race-ethnicity of the individuals arrested. This information may be elicited from the individual or inferred by the arresting officer. The LAPD recognizes the categories Asian, black, Latino, white and other, which combined constitute 97.7% of all arrests on average. Occasionally, other categories such as Filipino, Korean, and Pacific Islander appear within the data. Given the sometimes fraught history between the LAPD and communities of color [see 29, 39, 44] we focus on patterns in the arrest of black and Latino individuals and therefore report

results for these two groups and for white individuals.

3.1.5 Defining Crimes

We define crimes as those incidents reported to the LAPD that are classified by the LAPD crime coding system into 226 recognized crime types. In a typical year, the LAPD collects reports on approximately 180,000 crimes. Again we do not distinguish between different types of crimes.

4 Results

We now turn to a consideration of potential biases induced by predictive policing. We test three null hypotheses: (1) arrest of minority individuals did not differ between control and treatment conditions in test divisions; (2) arrest rates overall did not differ between control and treatment conditions in test divisions; (3) the rate of arrests per crime was unchanged across treatment and control conditions.

The LAPD experiment was designed to test for differences in predictive accuracy and impact on crime between control and treatment (predictive policing) conditions. Here we examine arrest patterns on control and treatment days (Table 1). Because in North Hollywood there was a lower number of treatment days ($n = 70$) compared to control days ($n = 82$) in the experiment, we adjust North Hollywood treatment counts to a rate per 82 days to be comparable to control. We conduct a Cochran-Mantel-Haenszel test [4] to examine whether ethnicity is independent of the treatment condition. The CMH test is a generalization of a chi-square test where the test is repeated across a strata, in this case several divisions. Here we do not find evidence to reject the null hypothesis that treatment arrests are independent of ethnicity, the CMH test p-value is 0.6957 (Table 1). The CMH test assumes that the treatment effect is homogeneous across the three divisions, we test this assumption using a Woolf test and do not find evidence to reject the null hypothesis of homogeneity (p-value 0.94337).

In Table 2 we test whether the total number of arrests were higher or lower on treatment days at the division level. Here we find that arrests were unchanged in Foothill and North Hollywood (using a chi square goodness of fit test for equal rates) and arrests were slightly lower in Southwest on treatment days (marginally significant at the 0.06 level).

We also examine arrest patterns in control and treatment boxes (Table 3). Recall that each participating division of the LAPD was allocated 20 prediction boxes, each 500x500 feet in size, each day. Control and treatment days were randomly determined. Again we test the association between treatment and arrest ethnicity using a CMH test. Here we do not find evidence to reject the null

hypothesis that treatment arrests in boxes are independent of ethnicity, the CMH test p-value is 0.3952 (Woolf test p-value = 0.720528).

In Table 4 we test whether arrests were higher or lower on treatment days within predicted boxes. Here we find a statistically significant increase in arrests (approximately double) in treatment boxes in all three divisions. To understand why arrests are higher in treatment boxes, in Table 5 we adjust the arrest rate to control for the higher overall rate of crime in treatment boxes. Our null hypothesis is that if predictive policing is “fair,” then the percentage of crimes in an area leading to an arrest should be equal across treatment and control conditions. Here we find this to be the case in North Hollywood (arrests per crime 7.0% control, 5.6% treatment, p-value 0.34) and Southwest (arrests per crime 10.3% control, 10.7% treatment, p-value 0.75). The arrest rate per crime is slightly lower in treatment boxes in Foothill (arrests per crime 14.9% control, 8.5% treatment, p-value 0.009).

5 Discussion and Conclusions

The stated goal of the analyses present above was to assess the degree to which arrest rates were impacted by the introduction of predictive policing in three divisions patrolled by the LAPD. Special attention was paid to arrest rates by the race-ethnicity of the individuals detained. Our null hypotheses were: (1) arrest of minority individuals did not differ between control and treatment conditions in test divisions; (2) arrest rates overall did not differ between control and treatment conditions in test divisions; (3) the rate of arrests per crime was unchanged across treatment and control conditions.

The evidence presented does not allow us to reject null hypothesis (1). There is no significant difference in the arrest proportions of minority individuals between treatment and control conditions. We also cannot reject hypothesis (2) at the division level. Arrests rates overall are the same on control and treatment days within the test divisions as a whole. However, we do reject null hypothesis (2) at the box level. Arrests were higher overall in treatment prediction boxes. We therefore tested hypothesis (3) to see if the higher arrest rate in treatment boxes is explained by an overall higher crime rate in treatment boxes. We fail to reject the null hypothesis (3). Arrest rates per crime do not differ across treatment and control conditions.

Clearly, arrests are a common part of day-to-day police operations. The introduction of predictive policing did not increase arrests overall, though treatment prediction boxes did see significantly more arrests than control prediction boxes. The increase arrests in treatment prediction boxes is perhaps understandable given that algorithmic crime predictions are more accurate than those produced by existing best practice [41].

The present study has several important limitations. Arrests are an imperfect proxy for other types of police contacts including stops, searches and detentions short of arrest. It is possible that predictive policing induced increases in these other categories of police contacts, without a concomitant impact on arrests. For this to hold true it would have to be the case that the rate of arrest actually declined as these other precursor contacts increased, leaving overall arrest numbers unchanged. This hypothetical downward adjustment in arrests would have to hold not only for the experimental deployment period overall, but also for randomly assigned treatment days. We do not have sufficient data to exclude such dynamics, but they seem improbable on the face of it.

Second, the analyses do not provide any guidance on whether arrests are themselves systemically biased. Such could be the case, for example, if black and Latino individuals experienced arrest at a rate disproportionate to their share of offending [50]. The current study is only able to ascertain that arrest rates for black and Latino individuals were not impacted, positively or negatively, by using predictive policing. Future research could seek to test whether the situational conditions surrounding arrests and final dispositions differ in the presence of predictive policing.

Finally, the results reported herein pertain for the narrowly-defined place-based predictive policing model used in the Los Angeles predictive policing experiment [41]. This model focused on reported crime data for a limited set of crimes including burglary, car theft and burglary from vehicle and used only information on crime location and time. Predictions were made for small 500x500 foot boxes and changed every day. Under those conditions we can conclude that predictive policing did not result in biased arrests. Whether the same outcomes would hold given changes in implementation is uncertain. If the exact same data types and methods are applied in a different location there may be reason to be optimistic. However, if the data types change, for example to focus on discretionary crimes or arrests, or if personal information is incorporated into predictions, then pessimism may be warranted. At the same time, we should ask whether police would be negligent if they had data or information that led to accurate forecasts of crime risk but failed to act on it for fear of potential bias [22]. Continued empirical scrutiny along with careful policy development will be needed to guard against bias in predictive policing and ensure fairness in outcomes.

6 Acknowledgements

This work was supported in part by US National Science Foundation Grants SCC-1737585, SES-1343123, ATD-1737996 and ATD-1737770. We thank the Los Angeles Police Department for providing data used in this research.

References

- [1] <https://www.aclu.org/other/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice>, 2016.
- [2] <https://www.sciencenews.org/blog/science-public/data-driven-crime-prediction-fails-erase-human-bias>, 2017.
- [3] <http://www.pbs.org/newshour/rundown/column-big-data-analysis-police-activity-inherently-biased/>, 2017.
- [4] Alan Agresti and Maria Kateri. Categorical data analysis. In *International encyclopedia of statistical science*, pages 206–208. Springer, 2011.
- [5] Frank R Baumgartner, Derek A Epp, Kelsey Shoub, and Bayard Love. Targeting young men of color for search and arrest during traffic stops: evidence from north carolina, 2002-2013. *Politics, Groups, and Identities*, pages 1–25, 2016.
- [6] Charlie Beck. Use of force year-end review executive summary 2015. Report, Los Angeles Police Department, 2016.
- [7] Katherine Beckett, Kris Nyrop, and Lori Pfingst. Race, drugs, and policing: Understanding disparities in drug delivery arrests. *Criminology*, 44(1):105–137, 2006.
- [8] Donald J. Black. Production of crime rates. *American Sociological Review*, 35(4):733–748, 1970.
- [9] Donald J. Black and Albert J. Reiss Jr. Police control of juveniles. *American Sociological Review*, 35(1):63–77, 1970.
- [10] Kate J Bowers, Shane D Johnson, and Ken Pease. Prospective hot-spotting the future of crime mapping? *British Journal of Criminology*, 44(5):641–658, 2004.
- [11] Anthony A Braga and Brenda J Bond. Policing crime and disorder hot spots: A randomized controlled trial. *Criminology*, 46(3):577–607, 2008.
- [12] Sarah Brayne. Big data surveillance: The case of policing. *American Sociological Review*, 82(5):977–1008, 2017.
- [13] James W Buehler. Racial/ethnic disparities in the use of lethal force by us police, 2010-2014. *American journal of public health*, (0):e1–e3, 2017.

- [14] R.V. Clarke and D. Weisburd. Diffusion of crime control benefits: Observations on the reverse of displacement. *Crime Prevention Studies*, 2:165–184, 1994.
- [15] Martin Degeling and Bettina Berendt. What is wrong about robocops as consultants? a technology-centric critique of predictive policing. *AI & Society*, 2017.
- [16] Ronnie A Dunn. Measuring racial disparities in traffic ticketing within large urban jurisdictions. *Public Performance & Management Review*, 32(4):537–561, 2009.
- [17] Charles R Epp, Steven Maynard-Moody, and Donald P Haider-Markel. *Pulled over: How police stops define race and citizenship*. University of Chicago Press, Chicago, IL, 2014.
- [18] Jeffrey Fagan, Anthony A Braga, Rod K Brunson, and April Pattavina. Stops and stares: Street stops, surveillance and race in the new policing. 2015.
- [19] Jeffrey Fagan, Amanda Geller, Garth Davies, and Valerie West. *Street Stops and Broken Windows Revisited: The Demography and Logic of Proactive Policing in a Safe and Changing City*. Race, Ethnicity and Policing. New York University, New York, NY, 2010.
- [20] Amy Farrell and Dean Jack McDevitt. Rhode island traffic stop statistics data collection study 2004-2005. Report, Institute on Race and Justice, April 2006 2006.
- [21] Andrew Guthrie Ferguson. Predictive policing and reasonable suspicion. *Emory LJ*, 62:259–324, 2012.
- [22] Andrew Guthrie Ferguson. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York University Press, New York, 2017.
- [23] Andrew Guthrie Ferguson. Policing predictive policing. *Washington University Law Review*, 94, in press.
- [24] William A Geller and Hans Toch. *And justice for all: Understanding and controlling police abuse of force*. Police Executive Research Forum Washington, DC, 1995.
- [25] Andrew Gelman, Jeffrey Fagan, and Alex Kiss. 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):813–823, 2007.

- [26] Sharad Goel, Justin M Rao, and Ravi Shroff. Precinct or prejudice? understanding racial disparities in new york city’s stop-and-frisk policy. *The Annals of Applied Statistics*, 10(1):365–394, 2016.
- [27] David A Harris. Frisking every suspect: The withering of terry. *UC Davis L. Rev.*, 28:1, 1994.
- [28] David A Harris. The stories, the statistics, and the law: Why driving while black matters. *Minn. L. Rev.*, 84:265, 1999.
- [29] Steven Kelly Herbert. *Policing space: Territoriality and the Los Angeles police department*. U of Minnesota Press, 1997.
- [30] William C Horrace and Shawn M Rohlin. How dark is dark? bright lights, big city, racial profiling. *Review of Economics and Statistics*, 98(2):226–232, 2016.
- [31] Brian Jordan Jefferson. Predictable policing: Predictive crime mapping and geographies of policing and race. *Annals of the American Association of Geographers*, 2017.
- [32] Shane D Johnson. *Prospective crime mapping in operational context: Final report*.
- [33] Tammy R. Kochel, David B. Wilson, and Stephen D. Mastrofski. Effect of suspect race on officers arrest decisions. *Criminology*, 49(2):473–512, 2011.
- [34] Christopher S. Koper. Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly*, 12(4):649–672, 1995.
- [35] Joscha Legewie. Racial profiling and use of force in police stops: How local events trigger periods of increased discrimination. *American Journal of Sociology*, 122(2):379–424, 2016.
- [36] Thomas A. Loughran, Raymond Paternoster, Alex R. Piquero, and Greg Pogarsky. On ambiguity in perceptions of risk: Implications for criminal decision making and deterrence. *Criminology*, 49(4):1029–1061, 2011.
- [37] Kristian Lum and William Isaac. To predict and serve? *Significance*, 13(5):14–19, 2016.
- [38] Mona Lynch, Marisa Omori, Aaron Roussell, and Matthew Valasik. Policing the progressive city: The racialized geography of drug law enforcement. *Theoretical Criminology*, 17(3):335–357, 2013.

- [39] Cid Martinez. *The Neighborhood Has Its Own Rules: Latinos and African Americans in South Los Angeles*. NYU Press, 2016.
- [40] Albert J Meehan and Michael C Ponder. Race and place: The ecology of racial profiling african american motorists. *Justice Quarterly*, 19(3):399–430, 2002.
- [41] George Mohler, Martin B. Short, Sean Malinowski, Mark Johnson, George E. Tita, Andrea L. Bertozzi, and P. Jeffrey Brantingham. Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, 110(512):1399–1411, 2015.
- [42] G.O. Mohler, M.B. Short, P.J. Brantingham, F.P. Schoenberg, and G.E. Tita. Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108, 2011.
- [43] Lyria Bennett Moses and Janet Chan. Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society*, 2016.
- [44] Ana Muniz. *Police, Power, and the Production of Racial Boundaries*. Rutgers University Press, 2015.
- [45] Justin Nix, Bradley A. Campbell, Edward H. Byers, and Geoffrey P. Alpert. A bird’s eye view of civilians killed by police in 2015. *Criminology & Public Policy*, 16(1):309–340, 2017.
- [46] Kenneth J Novak. Disparity and racial profiling in traffic enforcement. *Police Quarterly*, 7(1):65–96, 2004.
- [47] E Ashby Plant and B Michelle Peruche. The consequences of race for police officers’ responses to criminal suspects. *Psychological Science*, 16(3):180–183, 2005.
- [48] Greg Ridgeway. Assessing the effect of race bias in post-traffic stop outcomes using propensity scores. *Journal of quantitative criminology*, 22(1):1–29, 2006.
- [49] Victor M Rios. *Punished: Policing the lives of Black and Latino boys*. NYU Press, 2011.
- [50] Richard Rosenfeld and Robert Fornango. The impact of police stops on precinct robbery and burglary rates in new york city, 2003-2010. *Justice Quarterly*, 31(1):96–122, 2014.

- [51] Amie M Schuck. The masking of racial and ethnic disparity in police use of physical force: The effects of gender and custody status. *Journal of Criminal Justice*, 32(6):557–564, 2004.
- [52] Andrew D. Selbst. Disparate impact in big data policing. *Georgia Law Review*, 49:1–77, 2017.
- [53] Andrew D. Selbst. A mild defense of our new machine overlords. *Vanderbilt Law Review En Banc*, 70:87–104, 2017.
- [54] Lawrence W Sherman and John E Eck. *Policing for crime prevention*, pages 295–397. 2002.
- [55] Lawrence W Sherman, Patrick R Gartin, and Michael E Buerger. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27–56, 1989.
- [56] Lawrence W. Sherman and David Weisburd. General deterrent effects of police patrol in crime hot spots: A randomized, controlled trial. *Justice Quarterly*, 12:625–648, 1995.
- [57] Douglas A. Smith, Christy A. Visser, and Laura A. Davidson. Police control of juveniles. *Equity and discretionary justice: Influence of race on police arrest decisions*, 75(1):234–249, 1984.
- [58] Michael R Smith and Matthew Petrocelli. Racial profiling? a multivariate analysis of police traffic stop data. *Police Quarterly*, 4(1):4–27, 2001.
- [59] Forrest Stuart. *Down, out, and under arrest: Policing and everyday life in Skid Row*. University of Chicago Press, 2016.
- [60] Cody W Telep, David Weisburd, CharlotteE Gill, Zoe Vitter, and Doron Teichman. Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review. *Journal of Experimental Criminology*, 10(4):515–548, 2014.
- [61] Patricia Warren, Donald Tomaskovic-Devy, William Smith, Matthew Zin-graff, and Marcinda Mason. Driving while black: Bias processes and racial disparity in police stops. *Criminology*, 44(3):709–738, 2006.
- [62] David Weisburd. Place-based policing. *Ideas in American Policing*, 9:1–15, 2008.

- [63] David Weisburd and John E. Eck. What can police do to reduce crime, disorder, and fear? *The ANNALS of the American Academy of Political and Social Science*, 593(1):42–65, 2004.
- [64] David Weisburd, Laura A Wyckoff, Justin Ready, John E Eck, Joshua C Hinkle, and Frank Gajewski. Does crime just move around the corner? a controlled study of spatial displacement and diffusion of crime control benefits. *Criminology*, 44(3):549–592, 2006.
- [65] Brian R. Wyant, Ralph B. Taylor, Jerry H. Ratcliffe, and Jennifer Wood. Deterrence, firearm arrests, and subsequent shootings: A micro-level spatio-temporal analysis. *Justice Quarterly*, 29(4):524–545, 2012.

7 Tables and Figures

Table 1: Arrests vs. ethnicity on control and treatment days in three LAPD Divisions. † indicates counts adjusted for treatment days to a rate (per 82 days) to match control. Cochran-Mantel-Haenszel test $M^2 = 0.72581$, $df = 2$, p-value = 0.6957. Woolf test p-value = 0.94337

Division	Race/ Ethnicity	Control Arrests	Treatment Arrests
FH	Black	117	118
FH	Latino	700	756
FH	White	199	218
NH	Black	241	228†
NH	Latino	927	880†
NH	White	446	423†
SW	Black	1887	1818
SW	Latino	843	811
SW	White	110	73

Table 2: Total arrests on control vs treatment days aggregated by division. p-value is given for a chi square goodness of fit test for equal rates. † indicates counts adjusted for treatment days to a rate (per 82 days) to match control.

Division	Control Arrests	Treatment Arrests	p-value
FH	1016	1092	0.1
NH	1614	1531†	0.14
SW	2840	2702	0.06

Table 3: Arrests in control and treatment boxes for all three LAPD Divisions. † indicates counts adjusted for treatment days to a rate (per 82 days) to match control. Cochran-Mantel-Haenszel test $M^2 = 1.8567$, $df = 2$, p-value = 0.3952. Woolf test p-value = 0.720528

Division	Condition	Black	Latino	White
FH	Control	2	23	8
FH	Treatment	9	37	15
NH	Control	9	19	6
NH	Treatment	15†	31†	11†
SW	Control	71	25	6
SW	Treatment	158	44	6

Table 4: Arrests on control vs treatment days in boxes aggregated by division. p-value is given for a chi square goodness of fit test for equal rates. † indicates counts adjusted for treatment days to a rate (per 82 days) to match control.

Division	Control Arrests	Treatment Arrests	p-value
FH	33	61	0.004
NH	34	57†	0.015
SW	102	208	< .001

Table 5: Predictive policing arrests in boxes as a proportion of reported crimes in treatment vs. control boxes. p-value is given for a 2 sample proportions test. † indicates counts adjusted for treatment days to a rate (per 82 days) to match control.

Division	Control	Treatment	p-value
FH arrests/crimes	33/221 14.9%	61/714 8.5%	0.0085
NH arrests/crimes	34/485 7.0%	57/1021† 5.6%	0.3421
SW arrests/crimes	102/993 10.3%	208/1937 10.7%	0.7452

