



DATE DOWNLOADED: Tue Aug 31 13:23:25 2021

SOURCE: Content Downloaded from [HeinOnline](#)

Citations:

Bluebook 21st ed.

Rashida Richardson, Jason M. Schultz & Kate Crawford, Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice, 94 N.Y.U. L. REV. ONLINE 15 (2019).

ALWD 6th ed.

Richardson, R.; Schultz, J. M.; Crawford, K. ., Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice, 94 N.Y.U. L. Rev. Online 15 (2019).

APA 7th ed.

Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. New York University Law Review Online, 94, 15-55.

Chicago 17th ed.

Rashida Richardson; Jason M. Schultz; Kate Crawford, "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice," New York University Law Review Online 94 (2019): 15-55

McGill Guide 9th ed.

Rashida Richardson, Jason M Schultz & Kate Crawford, "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice" (2019) 94 NYU L Rev Online 15.

AGLC 4th ed.

Rashida Richardson, Jason M Schultz and Kate Crawford, 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice' (2019) 94 New York University Law Review Online 15.

MLA 8th ed.

Richardson, Rashida, et al. "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice." New York University Law Review Online, 94, 2019, p. 15-55. HeinOnline.

OSCOLA 4th ed.

Rashida Richardson and Jason M Schultz and Kate Crawford, 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice' (2019) 94 NYU L Rev Online 15

Provided by:

Available Through: University of Cambridge

-- Your use of this HeinOnline PDF indicates your acceptance of HeinOnline's Terms and Conditions of the license agreement available at

<https://heinonline.org/HOL/License>

-- The search text of this PDF is generated from uncorrected OCR text.

DIRTY DATA, BAD PREDICTIONS: HOW CIVIL RIGHTS VIOLATIONS IMPACT POLICE DATA, PREDICTIVE POLICING SYSTEMS, AND JUSTICE

RASHIDA RICHARDSON,* JASON M. SCHULTZ† & KATE CRAWFORD‡

Law enforcement agencies are increasingly using predictive policing systems to forecast criminal activity and allocate police resources. Yet in numerous jurisdictions, these systems are built on data produced during documented periods of flawed, racially biased, and sometimes unlawful practices and policies (“dirty policing”). These policing practices and policies shape the environment and the methodology by which data is created, which raises the risk of creating inaccurate, skewed, or systemically biased data (“dirty data”). If predictive policing systems are informed by such data, they cannot escape the legacies of the unlawful or biased policing practices that they are built on. Nor do current claims by predictive policing vendors provide sufficient assurances that their systems adequately mitigate or segregate this data.

In our research, we analyze thirteen jurisdictions that have used or developed predictive policing tools while under government commission investigations or federal court monitored settlements, consent decrees, or memoranda of agreement stemming from corrupt, racially biased, or otherwise illegal policing practices. In particular, we examine the link between unlawful and biased police practices and the data available to train or implement these systems. We highlight three case studies: (1) Chicago, an example of where dirty data was ingested directly into the city’s predictive system; (2) New Orleans, an example where the extensive evidence of dirty policing practices and recent litigation suggests an extremely high risk that dirty data was or could be used in predictive policing; and (3) Maricopa County, where despite extensive evidence of dirty policing practices, a lack of public transparency about the details of various predictive policing systems restricts a proper assessment of the risks. The implications of these findings have widespread ramifications for predictive policing writ large. Deploying predictive policing systems in jurisdictions with extensive histories of unlawful police practices presents elevated risks that dirty data will lead to flawed or unlawful predictions, which in turn risk perpetuating additional harm via feedback loops throughout the criminal justice system. The use of predictive policing must be treated with high levels of caution and mechanisms for the public to know, assess, and reject such systems are imperative.

INTRODUCTION16

* Director of Policy Research, AI Now Institute at New York University.

† Professor of Clinical Law, New York University; Law & Policy Area Lead, AI Now Institute at New York University.

‡ Distinguished Research Professor, New York University; Principal Researcher, Microsoft Research; Co-Founder, AI Now Institute at New York University. We wish to thank Andrew Ferguson, Barry Friedman, Meredith Whittaker, and Vincent Southerland for their invaluable insights and feedback on this essay. Copyright © 2019 by Rashida Richardson, Jason M. Schultz & Kate Crawford.

| | |
|---|----|
| A. <i>Dirty Data and “Juked Stats”</i> | 16 |
| B. <i>What Is Predictive Policing?</i> | 21 |
| I. THE RISK OF PREDICTIVE POLICING RELYING ON DIRTY DATA | 26 |
| A. <i>Overview of Study</i> | 26 |
| B. <i>Case Study 1: Chicago</i> | 28 |
| C. <i>Case Study 2: New Orleans</i> | 34 |
| D. <i>Case Study 3: Maricopa County</i> | 37 |
| II. THE PROBLEMS OF USING OF POLICE DATA AS CURRENTLY CONSTITUTED | 40 |
| A. <i>Confirmation Feedback Loop</i> | 40 |
| B. <i>The Boundless Risk of Biased Police Data</i> | 46 |
| CONCLUSION | 48 |
| APPENDIX | 50 |

INTRODUCTION

A. *Dirty Data and “Juked Stats”*

In the crime TV series *The Wire*, police are regularly instructed to shift arrest patterns and manipulate crime statistics in order to give the appearance of crime reduction. This is a practice called “juking the stats.” “If the felony rate doesn’t fall, you most certainly will,” one officer is instructed. The warning continues, “It’s Baltimore, gentlemen: the gods will not save you.”¹ Indeed, in Baltimore itself, the last ten years have resulted in federal investigations into systemic data manipulation, police corruption, falsifying police reports, and violence, including robbing residents, planting evidence, extortion, unconstitutional searches, and other corrupt practices that result in innocent people being sent to jail.² As of 2018, Baltimore faces as many as fifty-five potential lawsuits in connection to police corruption and tainted records.³

¹ *The Wire: Dead Soldiers* (HBO television broadcast Oct. 3, 2004).

² See CIVIL RIGHTS DIV., U.S. DEP’T OF JUSTICE, INVESTIGATION OF THE BALTIMORE CITY POLICE DEPARTMENT, 5–10 (2016), <https://www.justice.gov/crt/file/883296/download> (detailing the Department of Justice’s investigation of the Baltimore City Police Department, which found Baltimore policing practices in violation of both the Constitution and federal law); see also Luke Broadwater, *1,700 Cases Affected by Corrupt Baltimore Police Gun Trace Task Force, State’s Attorney Mosby Says*, BALT. SUN (June 5, 2018), <https://www.baltimoresun.com/news/maryland/crime/bs-md-ci-mosby-gttf-20180605-story.html> (reporting on the Baltimore City Police Department’s corrupt Gun Trace Task Force, a group of officers found guilty of robbing Baltimore residents).

³ Joy Lepola, *Baltimore Faces Up to 55 Possible Lawsuits Over Police Corruption*, FOX 45 NEWS (June 19, 2018), <https://foxbaltimore.com/features/operation-crime-justice/baltimore-faces-up-to-55-possible-lawsuits-over-police-corruption>; see also Justin Fenton, *‘Everything on the Table’ as Commission Begins Examining Corrupt Baltimore Police Gun Trace Task Force*, BALT.

While *The Wire* fictionalized Baltimore's very real history of police scandals, similar practices have been evidenced around the country. For example, in New York, "[m]ore than a hundred retired New York Police Department captains and higher-ranking officers said in a survey that the intense pressure to produce annual crime reductions led some supervisors and precinct commanders to manipulate crime statistics, according to two criminologists studying the department."⁴ These data manipulation practices were revealed in great detail through a survey and later a book published by two criminal justice professors, one of whom was a retired New York City Police Department (NYPD) captain, on the controversial crime statistics reporting program known as CompStat.⁵ The survey, book, and subsequent media reports revealed incidents of precinct commanders going to a crime scene to persuade victims to not file complaints in order to artificially reduce serious crime statistics, as well as officers engaging in corrupt practices like planting drugs on innocent people or falsifying records in order to meet arrests and summons quotas.⁶ The logic and goals of these practices were to keep reported numbers for "serious crimes" low, since that data had to be reported to the Federal Bureau of Investigation (FBI), while keeping the number of street arrests and summonses high, to give the appearance of community control over local crime.⁷

These "joked stats" were later confirmed by an NYPD-commissioned independent audit of the CompStat program,⁸ and the underlying practices were subsequently challenged in a series of lawsuits against NYPD. The

SUN (Oct. 16, 2018, 1:00 PM), <https://www.baltimoresun.com/news/maryland/crime/bs-md-ci-gttf-commission-first-meeting-20181016-story.html>.

⁴ William K. Rashbaum, *Retired Officers Raise Questions on Crime Data*, N.Y. TIMES (Feb. 6, 2010), <https://www.nytimes.com/2010/02/07/nyregion/07crime.html>.

⁵ See generally JOHN A. ETERNO & ELI B. SILVERMAN, *THE CRIME NUMBERS GAME: MANAGEMENT BY MANIPULATION* (2012).

⁶ See, e.g., John Marzulli, *We Fabricated Drug Charges Against Innocent People to Meet Arrest Quotas, Former Detective Testifies*, N.Y. DAILY NEWS (Oct. 13, 2011), <https://www.nydailynews.com/news/crime/fabricated-drug-charges-innocent-people-meet-arrest-quotas-detective-testifies-article-1.963021> (discussing a former NYPD narcotics detective's testimony about corrupt practices his squad engaged in to meet department quotas); Rashbaum, *supra* note 4 (detailing practice of persuading victims not to file complaints); Graham Rayman, *The NYPD Tapes: Inside Bed-Stuy's 81st Precinct*, VILLAGE VOICE (May 4, 2010), <https://www.villagevoice.com/2010/05/04/the-nypd-tapes-inside-bed-stuys-81st-precinct/> (discussing the content of hundreds of hours of recordings on the job by a Brooklyn police officer).

⁷ See Ryan Jacobs, *Just Like in "The Wire," Real FBI Crime Stats Are "Joked,"* MOTHER JONES (June 19, 2012, 9:49 PM), <https://www.motherjones.com/crime-justice/2012/06/fbi-crime-stats-fudged-the-wire-nypd/3/> (discussing the influence of FBI reporting requirements); Rashbaum, *supra* note 4 (discussing this data distortion).

⁸ See DAVID N. KELLEY & SHARON L. MCCARTHY, *THE REPORT OF THE CRIME REPORTING REVIEW COMMITTEE TO COMMISSIONER RAYMOND W. KELLY CONCERNING COMPSTAT AUDITING* 41–43 (2013), http://www.nyc.gov/html/nypd/downloads/pdf/public_information/crime_reporting_review_committee_final_report_2013.pdf (discussing the downgrading of reported crimes).

lawsuits resulted in a federal court opinion of almost two hundred pages finding that NYPD engaged in over a decade of unconstitutional and racially biased practices and policies, which required systemic reforms and monitoring by a federal court for compliance.⁹ Similarly, Baltimore's decades of police corruption resulted in a consent decree, signed in April 2017, between the City of Baltimore and the Department of Justice to address and reform police corruption and other unlawful practices.¹⁰ These court-based interventions primarily focus on cleaning up specific unconstitutional and corrupt processes and practices. But once the specified reforms are met there is little or no consideration of the need to address the police data that remains as an artifact of the prior unlawful conduct, and continues to shape predictive policing software going forward. So in this Article, we ask different questions: How does policing function as a data creation practice? What happens to predictive policing systems when police data records contain falsified crimes, planted evidence, racially biased arrests, and other actions that produce dirty data? In the absence of standardized data collection practices, how often do police departments or police technology vendors independently validate police records for accuracy or bias? How often might dirty data be included as the ground truth influencing predictive policing systems and other actors throughout the criminal justice system? What other forms of suspect or manipulated data might be ingested by predictive policing systems and how might this skew the predictions and subsequent recommendations? Can dirty data be remedied for subsequent use or is there a deeper and insurmountable problem derived from police practices and policies?

"Dirty data" is a term commonly used in the data mining research community to refer to "missing data, wrong data, and non-standard representations of the same data."¹¹ For the purposes of this paper, we are expanding the term "dirty data" to include a new category that reflects the culture of data production in policing. This new category includes data that is derived from or influenced by corrupt, biased, and unlawful practices, including data that has been intentionally manipulated or "joked," as well as data that is distorted by individual and societal biases. Dirty data—as we use the term here—also includes data generated from the arrest of innocent people who had evidence planted on them or were otherwise falsely accused, in addition to calls for service or incident reports that reflect false claims of criminal activity.¹² In addition, dirty data incorporates subsequent uses that

⁹ See *Floyd v. City of New York*, 959 F. Supp. 2d 540, 666 (S.D.N.Y. 2013).

¹⁰ See Consent Decree, *United States v. Police Dep't of Baltimore City*, No. 1:17-cv-00099-JKB (D. Md. Jan. 12, 2017), <https://www.justice.gov/opa/file/925056/download>.

¹¹ Won Kim et al., *A Taxonomy of Dirty Data*, 7 DATA MINING & KNOWLEDGE DISCOVERY 81, 81 (2003).

¹² See *infra* Section II.A.

further distort police records, such as the systemic manipulation of crime statistics to try to promote particular public relations, funding, or political outcomes. Importantly, data can be subject to multiple forms of manipulation at once, which makes it extremely difficult, if not impossible, for systems trained on this data to detect and separate “good” data from “bad” data, especially when the data production process itself is suspect. This challenge is notable considering that some prominent predictive policing experts assume that the problems of “dirty data” in policing can be isolated and repaired through classic mathematical, technological, or statistical techniques.¹³

For example, in 2015, NYPD entered a contract with Philadelphia-based technology company Azavea for a predictive policing system that would use NYPD’s historical crime data, among other factors, to predict where crime is likely to occur in the future in order to help precincts determine where to dispatch officers.¹⁴ In June 2018, the Baltimore Police Department (BPD) also expressed interest in acquiring a predictive policing system and other systems using police data.¹⁵ However, to date, neither NYPD, BPD, nor any of their technology vendors have clarified how they intend to address the “dirty data” problems that these systems may have, or even given assurances that all dirty data has been isolated and kept away from these systems. Given that these systems are shaped in large part by prior policing patterns, often reinforcing already known or ingrained biases, such risks are likely substantial.¹⁶ As an opinion piece in the *Baltimore Sun* noted,

¹³ See, e.g., P. Jeffrey Brantingham, *The Logic of Data Bias and Its Impact on Place-Based Predictive Policing*, 15 OHIO ST. J. CRIM L. 473, 485 (2018) (“The conclusion is that we need to work hard to figure out how to detect and correct for biases in police data rather than rejecting such data out of hand or accepting it without further thought.”).

¹⁴ See Laura Nahmias & Miranda Neubauer, *NYPD Testing Crime-Forecast Software*, POLITICO (July 8, 2015, 5:52 AM), <https://www.politico.com/states/new-york/city-hall/story/2015/07/nypd-testing-crime-forecast-software-090820>; N.Y.C. POLICE DEP’T, DEVELOPING THE NYPD’S INFORMATION TECHNOLOGY 7, <http://www.nyc.gov/html/nypd/html/home/POA/pdf/Technology.pdf> (“Drawing on historical data about past crime, including time, date, seasonal patterns, location, and crime type, the algorithms provide precinct and other commanders with an informed guide to emerging crime patterns and the deployment of resources within their respective areas of command.”).

¹⁵ See BALT. POLICE DEP’T, BALTIMORE POLICE DEPARTMENT TECHNOLOGY RESOURCE STUDY 61 (2018), https://www.baltimorepolice.org/sites/default/files/General%20Website%20PDFs/BPD_Final_Technology_Inventory_Study_06-21-18.pdf (explaining what technology resources the department needs to succeed); see also Caroline Haskins, *Predictive Policing Tool’s Website Exposes Login Pages for 17 US Police Departments*, VICE: MOTHERBOARD (Oct. 30, 2018, 1:33 PM), https://motherboard.vice.com/en_us/article/wj9v9q/predictive-policing-tools-website-exposes-login-pages-for-17-us-police-departments (describing Baltimore as a city with a PredPol login that suggests former or future use of the service even though there is no evidence of a finalized contract).

¹⁶ See generally MICHELLE ALEXANDER, *THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS* (2012) (arguing that racial discrimination in the criminal justice system, directed against Black and brown men, has led to a modern Jim Crow system); EMMA PIERSON ET

“Deploying officers based on crime statistics will simply return them to where they concentrate their time. As a result, the data often push officers into the same over-policed and over-criminalized communities.”¹⁷ This becomes part of what is known as the “bias in, bias out” concern regarding predictive systems.¹⁸

This problem is further intensified when policing data is tainted by corruption and other unconstitutional and unethical police practices. To evaluate these risks, we identified thirteen jurisdictions where there was public documentation showing an overlap in time between development or use of predictive policing systems and government investigations, consent decrees, or other documentation of corrupt, racially biased, or otherwise illegal police practices. We then compared the evidence from the government investigations and federal court adjudications with publicly available information regarding each jurisdiction’s use of predictive policing systems to determine whether dirty data was available to train or implement those systems during the periods of unlawful and biased police activity. We then looked to see if there was any publicly available evidence that might suggest an elevated risk of bias from use of dirty data.

Our analysis revealed nine jurisdictions where police data generated during periods when the department was found to have engaged in various forms of unlawful and biased police practices was available to train or otherwise inform predictive policing systems. In these jurisdictions, this overlap presented at least some risk that these predictive systems could be influenced by or in some cases perpetuate the illegal and biased police practices reflected in dirty data. While the relative lack of corporate and

AL., STANFORD COMPUTATIONAL POLICY LAB, A LARGE-SCALE ANALYSIS OF RACIAL DISPARITIES IN POLICE STOPS ACROSS THE UNITED STATES (2019), <https://5harad.com/papers/100M-stops.pdf> (finding racial disparities in police traffic stops and searches based on large-scale analysis of municipal and state police data); Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109, 1146–51 (2017) (“The assumptions behind predictive technologies are affected by unseen influences that may have unintended and discriminatory consequences.”); Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 WASH. U. L. REV. 35, 55–59 (2014) (analyzing uses of big data in policing and the Fourth Amendment concerns these uses provoke); Ezekiel Edwards, *Predictive Policing Software Is More Accurate at Predicting Policing than Predicting Crime*, HUFFPOST (Aug. 31, 2016, 2:58 PM), http://www.huffingtonpost.com/entry/predictive-policing-reform_us_57c6ffe0e4b0e60d31dc9120 (cautioning that the flaws of predictive policing software, like “racially skewed” historical crime data, may outweigh perceived benefits).

¹⁷ Michael Pinard, *Predicting More Biased Policing in Baltimore*, BALT. SUN (Apr. 10, 2018, 10:20 AM), <https://www.baltimoresun.com/news/opinion/oped/bs-ed-op-0411-predictive-policing-20180410-story.html>. Of course, if juked stats result in underreported crimes, the opposite effect of under-policing may occur. Both effects are symptoms of the bias problem, with over-policing communities of color raising far more serious civil rights issues.

¹⁸ See, e.g., Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. (manuscript at 3) (forthcoming 2019) (arguing that “a racially unequal past will necessarily produce racially unequal outputs”).

government transparency in many of these cases prevented further evaluation of the extent of any bias, we were able to observe several situations where there was a substantial risk. We also observed few, if any, efforts by police departments or predictive system vendors to adequately assess, mitigate, or provide assurances regarding the dirty data problem. Below, we highlight three of the nine case studies that show both direct and indirect linkages between dirty data and predictive policing as well as the dangers of obfuscating or inhibiting public oversight and accountability for how these systems are built and used.

B. What Is Predictive Policing?

Predictive policing generally describes any system that analyzes available data to predict either *where* a crime may occur in a given time window (place-based) or *who* will be involved in a crime as either victim or perpetrator (person-based). It is the latest iteration of data-driven crime analysis techniques that law enforcement agencies are increasingly relying on for crime control and forecasting.¹⁹ Few predictive policing vendors are fully transparent about how their systems operate, what specific data is used in each jurisdiction that deploys the technology, or what accountability measures the vendor employs in each jurisdiction to address potential inaccuracy, bias, or evidence of misconduct. Despite these looming questions, one known fact is that historical police data is the primary data source used to inform these systems, and, while the specific data categories will vary by system, the data can include information on past crimes (type of crime, time, and location), arrests and calls for service.²⁰ Some vendors exclude data that more obviously reflects biased and discretionary police practices, like arrest and stop data, but there is much less transparency about how vendors deal with categories of data where the embedded bias is less apparent, like call for service data.²¹

¹⁹ See Ferguson, *supra* note 16, at 1123–43 (describing the evolution of predictive policing); Joh, *supra* note 16, at 44–46 (describing geographic and individual-based predictive policing models). There is currently no definitive or chronological study on predictive policing. Studies have looked at data-driven policing, and other books and papers have focused on predictive policing, but nothing yet has bridged the gap. See Sarah Brayne, *Big Data Surveillance: The Case of Policing*, 82 AM. SOC. REV. 977, 982 (2017) (“[A]lthough there is strong theoretical work in surveillance studies, how big data surveillance plays out on the ground remains largely an open empirical question.”).

²⁰ See Andrew Guthrie Ferguson, *The Police Are Using Computer Algorithms to Tell if You’re a Threat*, TIME (Oct. 3, 2017), <http://time.com/4966125/police-departments-algorithms-chicago/>; see, e.g., AZAVEA, HUNCHLAB: UNDER THE HOOD 19 (2015), <https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf> (illustrating one company’s explanation of how it can make police forces more effective by using historical crime data and providing additional temporal and geographic data sets).

²¹ See, e.g., Robert Cheetham, *Why We Sold HunchLab*, AZAVEA: BLOG (Jan. 23, 2019), <https://www.azavea.com/blog/2019/01/23/why-we-sold-hunchlab/> (“We would not use data about

Though many may assume that police data is objective, it is embedded with political, social, and other biases.²² Indeed, police data is a reflection of the department's practices and priorities; local, state or federal interests; and institutional and individual biases.²³ In fact, even calling this information "data" could be considered a misnomer, since "data" implies some type of consistent scientific measurement or approach.²⁴ In reality there are no standardized procedures or methods for the collection, evaluation, and use of information captured during the course of law enforcement activities, and police practices are fundamentally disconnected from democratic controls, such as transparency and oversight.²⁵ This lack of rigorous methodology and accountability leaves significant room for subjectivity in how or what officers choose to record about their activities, and few incentives for police leadership or other government actors to interrogate or validate records for accuracy, bias, or misconduct, or to identify proactive reforms.²⁶

Moreover, there is no evidence of predictive policing vendors independently validating the police records within the jurisdictions where the technology is deployed. Instead, even those vendors willing to acknowledge biased data as a problem merely attempt to isolate or segregate it from what is presumably "clean" data instead of seeing it as an indicator of the potential unreliability of the entire data set from that jurisdiction. For

people—no arrests, no social media, no gang status, no criminal background information.”); PREDPOL, <https://www.predpol.com/law-enforcement/> (last visited Mar. 26, 2019) (“PredPol uses ONLY 3 data points—crime type, crime location, and crime date/time—to create its predictions.”).

²² See Brayne, *supra* note 19, at 997–1000 (arguing that police data is embedded with numerous biases); Brian Jordan Jefferson, *Predictable Policing: Predictive Crime Mapping and Geographies of Policing and Race*, 108 ANNALS AM. ASS'N GEOGRAPHERS 1, 2 (2018) (arguing that relying on official crime statistics further legitimizes racialized policing). See generally ANDREW GUTHRIE FERGUSON, *THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT* (2017) (analyzing “black data” and focusing on racial bias and the opaque nature of police data as dangers of predictive policing systems).

²³ See Barbara E. Armacost, *Organizational Culture and Police Misconduct*, 72 GEO. WASH. L. REV. 453, 494 (2004) (describing how the organizational culture in police departments affects misconduct and other unlawful practices); FERGUSON, *supra* note 22, at 143–66; see also Kristian Lum & William Isaac, *To Predict and Serve?*, SIGNIFICANCE, Oct. 2016, at 15–16 (analyzing the social consequences of relying on predictive policing systems that use biased data).

²⁴ See LISA GITELMAN & VIRGINIA JACKSON, *Introduction to “RAW DATA” IS AN OXYMORON* 1, 3, 1–12 (Lisa Gitelman ed., 2013) (“Data need to be imagined as data to exist and function as such, and the imagination of data entails an interpretive base.”).

²⁵ See Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. REV. 1827, 1855–65 (2015) (explaining why policing lacks administrative and congressional oversight and transparency); see also Marvin E. Wolfgang, *Uniform Crime Reports: A Critical Appraisal*, 111 U. PA. L. REV. 708, 713–17 (1963) (detailing the flaws in the FBI's Uniform Crime Reporting, reports that ideally capture statistics about policing and prosecution).

²⁶ Rachel Harmon, *Why Do We (Still) Lack Data on Policing?*, 96 MARQ. L. REV. 1119, 1129–33 (2013) (describing how cities and police departments actively inhibit the collection of information about police and how federal efforts of data collection and review are stymied by institutional and legal limitations).

example, the place-based predictive policing company PredPol claims that it primarily collects and analyzes so-called “victim data”—“time of day, location of crime and type of crime from reports.”²⁷ This is an attempt to position such data as objective and untainted. On the other hand, PredPol “excludes drug-related offenses”—which have known and well-documented racial disparities—“and traffic citation data from its predictions to remove officer bias from the equation and eliminate the risk of generating predictions based on officer discretion.”²⁸ Yet drug-related offenses and traffic stops are hardly the only crimes where officers exercise bias or discretion, let alone the only crimes where corruption, discrimination, racial profiling, or other dirty policing practices exist. For example, even deciding which circumstances to investigate, define, and document as a “crime” can be a matter of officer discretion, as our case studies show.

Moreover, examining the context and use of policing practices to generate data is also important because encounters with police are the most common point of entry for individuals into the criminal justice system. Many subsequent decision-making processes in criminal justice, including those during prosecution, pretrial services, adjudication, sentencing, parole, and corrections, derive their analysis from policing data inputs.²⁹ As such, an examination of policing practices and the data that acts as a record of such practices is required for any informed discussion of the possible harms and benefits of predictive policing systems whose conclusions are shaped by such data. Yet, despite this need for scrutiny, policing is often the least regulated of all the government agencies.³⁰ A driver of this imbalance is that law enforcement agencies are given undue deference by all branches of government.³¹

²⁷ See *Machine Learning and Policing*, PREDPOL: PREDICTIVE POLICING BLOG (July 19, 2017, 10:00 AM), <http://blog.predpol.com/machine-learning-and-policing> (explaining PredPol’s methodology and lauding its benefits).

²⁸ *Id.*

²⁹ See JENNIFER A. TALLON ET AL., CTR. FOR COURT INNOVATION, THE INTELLIGENCE-DRIVEN PROSECUTION MODEL: A CASE STUDY IN THE NEW YORK COUNTY DISTRICT ATTORNEY’S OFFICE, at vi–vii (2016), https://www.courtinnovation.org/sites/default/files/documents/IDPM_Research_Report_FINAL.PDF (describing the New York County District Attorney’s Office’s use of police data to improve prosecutorial decision-making); see also MHS, LEVEL OF SERVICE/CASE MANAGEMENT INVENTORY (2004), https://www.assessments.com/assessments_documentation/LSCMI_Tech_Brochure.pdf (describing the use of police data in parole case management); NORTHPOINTE, PRACTITIONERS GUIDE TO COMPAS 1–2 (2012), http://www.northpointeinc.com/files/technical_documents/FieldGuide2_081412.pdf (describing how Northpointe’s web-based tool uses police data to inform pre-trial, sentencing, parole, and probation decisions).

³⁰ Friedman & Ponomarenko, *supra* note 25, at 1843–48 (describing how police remain the least regulated by conventional checks and balances of government).

³¹ See *id.* (explaining how policing largely escapes legislative and administrative oversight).

It is a common fallacy that police data is objective and reflects actual criminal behavior, patterns, or other indicators of concern to public safety in a given jurisdiction. In reality, police data reflects the practices, policies, biases, and political and financial accounting needs of a given department.³² For instance, some data relevant to crime patterns and public safety indicators, such as police misconduct, are not reflected in police data or available publicly, which reflects underlying forms of political accounting and public relations.³³ Hence, actual crime data is often incomplete or distorted. The Department of Justice has estimated that less than half of violent crimes and even fewer household property crimes are reported to the police.³⁴ The type of criminal activity recorded by police also depends on which law enforcement agency has jurisdiction over which crimes.³⁵ Research also suggests that groups that feel less favorable toward local police are less likely to report crime they witness.³⁶ Even when reported, errors and bias in how the police record reported crimes result in distorted data. The Los Angeles Police Department (LAPD), for example, misrecorded a staggering 14,000 serious assaults as minor offenses from 2005 to 2012. This error was not discovered until 2015, by which time LAPD had already begun its work with the predictive policing company PredPol, though there is no evidence to confirm whether this erroneous data was used in the system.³⁷

While several prominent predictive policing vendors have acknowledged concerns about the inclusion of biased data in their systems, most vendors fail to account for these structural and systemic errors in the

³² See Rachel A. Harmon, *Promoting Civil Rights Through Proactive Policing Reform*, 62 STAN. L. REV. 1, 5 (2009) (discussing the lack of uniformity in police data reporting standards); see also Armacost, *supra* note 23, at 474 (discussing the shortcomings of police data records).

³³ See Armacost, *supra* note 23, at 474 (suggesting that police misconduct data is inaccessible either because of poor record keeping or deliberate hiding by law enforcement to prevent the data's use in litigation); Harmon, *supra* note 32 (discussing lack of police misconduct data).

³⁴ LYNN LANGTON ET AL., BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 238536, VICTIMIZATIONS NOT REPORTED TO THE POLICE, 2006–2010, at 4, tbl.1 (2012), <https://www.bjs.gov/content/pub/pdf/vnnp0610.pdf>.

³⁵ For example, the White Collar Crime Risk Zones Map satirically shows what predictive policing might look like if it were based on the data in the Financial Industry Regulatory Authority database of financial crimes instead of the NYPD database. See Brian Clifton et al., *White Collar Crime Risk Zones*, NEW INQUIRY MAG. (Mar. 2017), <https://whitecollar.thenewinquiry.com>.

³⁶ See EMILY EKINS, CATO INST., POLICING IN AMERICA: UNDERSTANDING PUBLIC ATTITUDES TOWARD THE POLICE. RESULTS FROM A NATIONAL SURVEY. 14 (2016) (discussing the results, suggesting this, of a joint national survey on policing between the Cato Institute and YouGov); see also Lum & Isaac, *supra* note 23, at 15 (“Bias in police records can also be attributed to levels of community trust in police, and the desired amount of local policing These effects manifest as unequal crime reporting rates throughout a precinct.”).

³⁷ Ben Poston et al., *LAPD Underreported Serious Assaults, Skewing Crime Stats for 8 Years*, L.A. TIMES (Oct. 15, 2015, 4:47 AM), <https://www.latimes.com/local/cityhall/la-me-crime-stats-20151015-story.html> (reporting on the misclassification of the crimes, which concealed a seven percent increase in violent crime in Los Angeles over that period).

data, often overestimating what can be remedied.³⁸ Not only is the challenge of identifying and correcting these problems difficult, if not insurmountable, but it also raises significant doubts about the ability to distinguish known problematic data categories, such as drug-related arrest data, from data categories that are customarily considered objective, such as calls for service data.³⁹ Moreover, even where such distinctions are possible, they would have to occur on a jurisdiction-by-jurisdiction basis, since police data collection and classification practices vary by department and are often performed in ways that make aggregate or comparative analysis impossible.⁴⁰ There is a dearth of objective and comprehensive analysis of the efficacy and impact of certain police practices or policies, coupled with an unwillingness or inability on the part of police departments to investigate and monitor themselves impartially.⁴¹ To overcome the challenge that this information deficit poses, our study relied on the findings of government-commissioned investigations, federal court monitored settlements, consent decrees, or memoranda of agreement. These types of agreements “begin with investigations of allegations of systemic police misconduct and, when the allegations are substantiated, end with comprehensive agreements designed to support constitutional and effective policing and restore trust between police and communities.”⁴² Additionally, the findings related to these investigations, agreements, and federal court adjudications are generally considered substantiated, despite the fact that the settlements involve no finding of guilt. The required reforms typically prohibit the identified

³⁸ For example, HunchLab’s guide to the place-based predictive policing system notes that bias in the enforcement of certain crimes can distort police data used in its system. As a corrective measure, it states that “[f]or quality of life type crimes, we tend to use records that reflect the public’s call for services, which does not suffer from an enforcement bias.” Yet the company fails to acknowledge or account for the fact that the public’s call for services can reflect other societal bias that can also distort the data, as discussed below in Section II.A. See HUNCHLAB, A CITIZEN’S GUIDE TO HUNCHLAB 26 (2017), <http://robertbrauneis.net/algorithms/HunchLabACitizensGuide.pdf>.

³⁹ See Harmon, *supra* note 26, at 1129–33 (“Though officers will collect information when police chiefs and local governments require them to do so, they will collect only that information and only in the form mandated. . . . Other times, police departments simply fail to produce records that could improve political and regulatory decision-making about intrusive police activities.”); see also NAT’L ACADS. OF SCI., ENG’G & MED., PROACTIVE POLICING: EFFECTS ON CRIME AND COMMUNITIES 292 (David Weisburd & Malay K. Majmudar eds., 2018) (noting that whether calls for service are racially biased is an open question).

⁴⁰ Harmon, *supra* note 26, at 1129 (“Even when departments collect information, they may do so in ways that make it impossible to aggregate the records or compare them with data from other departments.”).

⁴¹ See *id.* at 1130 (“But the reality is that public actors who shape policing—from the officers themselves to local politicians—often face incentives that undermine data collection and research on policing as well as distribution of information about policing to the public.”).

⁴² CIVIL RIGHTS DIV., U.S. DEP’T OF JUSTICE, THE CIVIL RIGHTS DIVISION’S PATTERN AND PRACTICE POLICE REFORM WORK: 1994-PRESENT 1 (2017), <https://www.justice.gov/crt/file/922421/download>.

problematic practices, and a federal judge or third party monitor has authority to enforce compliance with such prohibitions.⁴³ Given this, we conclude that the data generated during the time periods covered by these findings sufficiently reflects bias and misassumptions, at the very least, embedded within police practices.

There is significant research and litigation raising concerns of bias in policing and the broader criminal justice system, but much of this scrutiny is focused on specific actors, practices, or newly adopted systems.⁴⁴ Complimenting these approaches, we examine how individual and collective practices by actors in and outside of the criminal justice system are reflected in the data that is generated and subsequently used throughout the criminal justice system, often without adequate transparency, accountability, oversight, or public engagement. We argue that when dirty data exists in the criminal justice system, it is often systemic and pervasive; therefore, strategies to isolate or mitigate its impact on predictive systems, especially technological ones, are unlikely to eliminate the dirty data problem or rebut the presumption that other policing data from the same jurisdiction is unproblematic.

I

THE RISK OF PREDICTIVE POLICING RELYING ON DIRTY DATA

A. *Overview of Study*

For this Article, we identified thirteen jurisdictions where publicly available information showed an overlap in time between development or use of predictive policing systems and the existence of government-commissioned investigations, federal court monitored settlements, consent decrees, or memoranda of agreement that found that the police departments engaged in corrupt, racially biased, or otherwise illegal police practices. These jurisdictions include Baltimore, Maryland; Boston, Massachusetts;

⁴³ See *id.* at 20–25 (describing the typical structure of these reforms); see also Peter M. Shane, *Federal Policy Making by Consent Decree: An Analysis of Agency and Judicial Discretion*, 1987 U. CHI. LEGAL F. 241 (discussing the prevalence of implicit racial bias in sentencing).

⁴⁴ See, e.g., Mark W. Bennett, *The Implicit Racial Bias in Sentencing: The Next Frontier*, 126 YALE L.J. FORUM 391, 396 (2017); FERGUSON, *supra* note 22, at 3 (critiquing the rise of predictive policing as creating “black data” that disproportionately harms communities of color); NOAH ZATZ ET AL., UCLA LABOR CTR., GET TO WORK OR GO TO JAIL: WORKPLACE RIGHTS UNDER THREAT 4–6 (2016), <https://www.labor.ucla.edu/publication/get-to-work-or-go-to-jail/> (describing how probation and parole conditions force low-income people to choose between bad or potentially unpaid jobs and jail time); see also DAVID ROGERS, ACLU OF OR., ROADBLOCKS TO REFORM: DISTRICT ATTORNEYS, ELECTIONS, AND THE CRIMINAL JUSTICE STATUS QUO 4 (2016), http://aclu-or.org/sites/default/files/Roadblocks_to_Reform_Report_ACLUOR.pdf (detailing how district attorneys play an important role in blocking progressive criminal justice reform and maintain the status quo out of self-interest).

Chicago, Illinois; Ferguson, Missouri; Miami, Florida; Maricopa County, Arizona; Milwaukee, Wisconsin; New Orleans, Louisiana; New York, New York; Newark, New Jersey; Philadelphia, Pennsylvania; Seattle, Washington; and Suffolk County, New York.⁴⁵ We compared the substantiated evidence and other findings of unlawful or biased police practices from the Department of Justice investigations or federal court adjudications with publicly available information regarding the jurisdiction's predictive policing activities to determine whether the police data available to train or implement the predictive policing system(s) was generated during the periods of unlawful and biased police activity.

We also looked at publicly available information about the nature of the systems and the exact data and training models they used. However, the general lack of public transparency concerning policing and predictive policing systems often makes it difficult to draw a straight line between the dirty data produced by the police departments in the reviewed jurisdictions and the predictive policing systems deployed. Thus, as an initial finding, we identified nine jurisdictions where dirty data was available to train or inform predictive policing systems and four other jurisdictions where our research was not dispositive. We then looked more closely for publicly available evidence demonstrating either a direct or indirect link between dirty data and risks of predictive policing bias.

The following three case studies highlight our findings and conclusions: Chicago is an example of a jurisdiction where we found strong evidence to suggest that the predictive policing system was using dirty data. Second, New Orleans is an example of a jurisdiction where the extensive dirty policing practices and recent litigation suggest an extremely high likelihood that some dirty data was used with predictive policing, although because the public has been blocked from proper transparency and accountability mechanisms, the extent of the problem is not fully known. Finally, Maricopa County is an example of a jurisdiction where extensive dirty policing practices suggest an extremely high risk that any predictive policing application will end up using dirty data not only for the County itself, but also for adjacent jurisdictions where data or police resources are shared. Again, lack of public transparency and accountability inhibits a forensic examination of the risks. Maricopa County also demonstrates an emerging trend of local police departments engaging in immigration enforcement, which complicates many of the aforementioned issues regarding the constitutionality, transparency, accountability, and oversight of police practices.

⁴⁵ See *infra* Appendix.

B. Case Study 1: Chicago

The Chicago Police Department (CPD) has a lengthy and well-documented history of corrupt, abusive, and biased practices, dating back to a 1972 blue-ribbon panel finding of extreme police misconduct that disproportionately affected residents of color.⁴⁶ In the years since, there have been several notable investigations and legal challenges, including evidence of over one hundred cases of CPD torturing Black men between 1972 and 1991⁴⁷ and a lawsuit challenging CPD's inequitable deployment of police to emergency calls in neighborhoods with higher minority populations.⁴⁸ Given this breadth of issues, the remainder of our discussion of CPD's practices, policies, and data is limited to the last decade.

In March 2015, the ACLU of Illinois issued a groundbreaking report detailing CPD's fraught history of stop and frisk practices.⁴⁹ Using CPD records on stops that occurred in 2012 and 2013 and four months of contact card data⁵⁰ from 2014, the report concluded that a significant number of CPD stop and frisks were unlawful, and Black residents were disproportionately subjected to these unlawful stops.⁵¹ The report, which also found significant deficiencies in CPD data and data collection practices,⁵² led to a settlement agreement in August 2015 requiring ongoing independent evaluation of CPD practices and procedures, data collection, officer training, and reform of

⁴⁶ See BLUE RIBBON PANEL, THE MISUSE OF POLICE AUTHORITY IN CHICAGO 29 (1972), <https://chicagopatf.org/wp-content/uploads/2016/01/metcalfe-report-1972.pdf>.

⁴⁷ See *The Chicago Torture Archive*, POZEN FAMILY CTR. FOR HUMAN RIGHTS, UNIV. OF CHI., <http://web.archive.org/web/20161028161512/http://chicagotorturearchive.uchicago.edu/> (last visited Feb. 13, 2019) (archiving the court proceedings for victims of police torture); see also, e.g., *United States v. Burge*, 711 F.3d 803, 807 (7th Cir. 2013) ("Jon Burge joined the CPD in 1970 and rose to commanding officer of the violent crimes section in the 1980s, but his career was marked by accusations from over one hundred individuals who claimed that he and officers under his command tortured suspects . . . throughout the 1970s and 1980s."); *Hinton v. Uchtman*, 395 F.3d 810, 822 (7th Cir. 2005) (Wood, J., concurring) ("[A] mountain of evidence indicates that torture was an ordinary occurrence at the Area Two station of the Chicago Police Department.").

⁴⁸ See Complaint at 1, *Central Austin Neighborhood Ass'n v. City of Chicago*, 1 N.E.3d 967 (Ill. App. Ct. 2011) (No. 11-CH-37299), https://www.aclu-il.org/sites/default/files/field_documents/complaint.pdf (challenging the City of Chicago's failure to ensure the equitable deployment of police officers across the city, which resulted in delayed police responses and high rates of serious violent crimes in neighborhoods with significant ethnic minority populations).

⁴⁹ See ACLU OF ILL., STOP AND FRISK IN CHICAGO 5–6 (2015), https://www.aclu-il.org/sites/default/files/wp-content/uploads/2015/03/ACLU_StopandFrisk_6.pdf ("A review of how Terry stops have been used in Chicago demonstrates a persistent problem—inadequate training, supervision and monitoring of law enforcement in minority communities.").

⁵⁰ These are forms CPD officers fill out after a street stop.

⁵¹ See ACLU OF ILL., *supra* note 49, at 6–11.

⁵² *Id.* at 14–15 ("[T]he Chicago Police Department does not record stop and frisks in a way that reflects the full picture of what is happening on the streets of Chicago. Chicago does not have a single database of all stops available to the public and officers make no records of frisks.").

investigatory street stop practices.⁵³ A former federal judge oversaw the agreement and published regular public reports assessing CPD's compliance with the agreement's requirements, which revealed CPD's continued engagement in unlawful practices and the data reflected significant race and gender bias.⁵⁴

During this same period, Chicago received national attention due to public outcry and city-wide protests following the release of a videotape showing the fatal shooting of Laquan McDonald, a Black 17-year-old, by a CPD officer.⁵⁵ This led to the Illinois Attorney General's December 2015 request to the Department of Justice to investigate CPD, which resulted in a yearlong investigation of CPD and the Independent Police Review Authority, the body responsible for investigating police misconduct. For this investigation, federal officials reviewed CPD records between 2011 and 2016, performed local visits, and met with community members, City officials, CPD staff, and local unions.⁵⁶ A report on this investigation concluded that CPD engaged in a pattern or practice of unconstitutional use of force; poor data collection to identify and address unlawful conduct; systemic deficiencies in training and supervision; systemic deficiencies in accountability systems that contribute to the pattern or practice of unconstitutional conduct; and unconstitutional conduct that disproportionately affects Black and Latino residents.⁵⁷

⁵³ See INVESTIGATORY STOP AND PROTECTIVE PAT DOWN SETTLEMENT AGREEMENT 2–8 (2015), <https://www.aclu-il.org/sites/default/files/wp-content/uploads/2015/08/2015-08-06-Investigatory-Stop-and-Protective-Pat-Down-Settlement-Agreement....pdf>.

⁵⁴ See ARLANDER KEYS, THE CONSULTANT'S SECOND SEMI-ANNUAL REPORT: INVESTIGATORY STOP & PROTECTIVE PAT DOWN AGREEMENT 171–72 (2018), https://www.aclu-il.org/sites/default/files/field_documents/the_consultants_second_semi-annual_report.pdf; see also ARLANDER KEYS, THE CONSULTANT'S FIRST SEMI-ANNUAL REPORT ON THE INVESTIGATORY STOP AND PROTECTIVE PAT DOWN AGREEMENT FOR THE PERIOD JANUARY 1, 2016 – JUNE 30, 2016, at 165–68 (2017), <https://www.aclu-il.org/sites/default/files/wysiwyg/the-consultants-first-semiannual-report-3-23-17.pdf>.

⁵⁵ See Monica Davey & Mitch Smith, *Chicago Protests Mostly Peaceful After Video of Police Shooting Is Released*, N.Y. TIMES (Nov. 24, 2015), <https://www.nytimes.com/2015/11/25/us/chicago-officer-charged-in-death-of-black-teenager-official-says.html>.

⁵⁶ See CIVIL RIGHTS DIV., U.S. DEP'T OF JUSTICE & U.S. ATTORNEY'S OFFICE FOR THE N. DIST. OF ILL., INVESTIGATION OF THE CHICAGO POLICE DEPARTMENT 1–2 (2017), <https://www.justice.gov/opa/file/925846/download>. In the same month that he requested the investigation, Chicago Mayor Rahm Emanuel separately created a Police Accountability Task Force, which subsequently released a report with more than 100 recommendations for reform in April 2016. See POLICE ACCOUNTABILITY TASK FORCE, RECOMMENDATIONS FOR REFORM: RESTORING TRUST BETWEEN THE CHICAGO POLICE AND THE COMMUNITIES THEY SERVE 22 (2016), <https://chicagopatf.org/wp-content/uploads/2016/04/PATF-Complete-Recommendations.pdf>.

⁵⁷ See *id.* at 145. The report described video or other evidence demonstrating CPD's unconstitutional conduct that was falsely or inaccurately described in reports provided by officers. In some instances, these inaccurate reports resulted in false arrests and convictions. The report

Despite these findings, the Department of Justice, under direction of then-Attorney General Jeff Sessions, announced that it would not seek a consent decree to reform CPD.⁵⁸ This prompted several lawsuits against the City of Chicago to seek a consent decree to reform CPD that would address the findings and recommendations of the Department of Justice investigation.⁵⁹ The City of Chicago opted to negotiate a consent decree with only one plaintiff of the lawsuits, the Illinois Attorney General, and it entered a Memorandum of Agreement with the other plaintiffs, over a dozen community and civil rights organizations, agreeing to pause the lawsuits during the consent decree negotiations and providing the local organizations the right to object if the decree is inadequate.⁶⁰ In September 2018, the Illinois Attorney General's Office and the City filed a draft consent decree in federal court, and in late January U.S. District Judge Robert Dow, Jr. approved the plan.⁶¹ Though the consent decree includes extensive reforms of CPD practices, policies, and oversight, the community organization plaintiffs have identified several deficiencies.⁶²

Amid these years of overlapping investigations of and challenges to

concluded that the extent of CPD's unconstitutional conduct may be larger than CPD records indicate, given the large volume of reported incidents that do not have exculpatory or corroborating evidence. This further demonstrates our concerns regarding "dirty data" generated by unlawful and biased police practices. *Id.* at 36–37, 74–79.

⁵⁸ United States' Statement of Interest Opposing Proposed Consent Decree at 3–4, *Illinois v. City of Chicago*, No. 17-cv-6260, 2019 WL 398703 (N.D. Ill. Oct. 12, 2018), <https://www.documentcloud.org/documents/5001703-Trump-Chicago-police-reform.html>; see also Dan Hinkel, *U.S. Attorney General Says Chicago Police Consent Decree Should Be Tossed, While Activists Seek Tighter Rules*, CHI. TRIB. (Oct. 12, 2018, 6:50 PM), <https://www.chicagotribune.com/news/local/breaking/ct-met-consent-decree-chicago-police-shootings-20181012-story.html>.

⁵⁹ See, e.g., Amended Complaint, *Cmtys. United v. City of Chicago*, No. 17-cv-7151 (N.D. Ill. Nov. 28, 2017), https://www.aclu-il.org/sites/default/files/field_documents/amended_complaint_0.pdf; Complaint, *Illinois v. City of Chicago*, 2019 WL 398703 (N.D. Ill. 2019) (No. 17-cv-6260), http://www.illinoisattorneygeneral.gov/pressroom/2017_08/City_of_Chicago_lawsuit_filedcomplaint.pdf.

⁶⁰ See Memorandum of Agreement Between the Office of the Illinois Attorney General and the City of Chicago and *Campbell v. City of Chicago* Plaintiffs and *Communities United v. City of Chicago* Plaintiffs (Mar. 16, 2018), https://www.aclu-il.org/sites/default/files/field_documents/executed_moa.pdf [hereinafter Chicago Memorandum of Agreement].

⁶¹ See Consent Decree, *Illinois v. City of Chicago*, No. 17-cv-6260 (N.D. Ill. 2019), <http://chicagopoliceconsentdecree.org/wp-content/uploads/2018/09/Illinois-v.-Chicago-Final-Consent-Decree-with-signatures.pdf>; Dan Hinkel, *Judge Approves Historic Court Order Aimed at Reforming Chicago Police Department: 'Let Us Begin,'* CHI. TRIB. (Feb. 1, 2019), <https://www.chicagotribune.com/news/local/breaking/ct-met-chicago-police-oversight-decree-20190131-story.html>.

⁶² See *CAMPBELL v. CITY OF CHICAGO* PLAINTIFFS, TRANSFORMING THE CPD: CONSENT DECREE REQUIREMENTS, https://www.aclu-il.org/sites/default/files/field_documents/campbell_v._city_of_chicago_-_transforming_the_cpd.pdf (detailing these shortcomings).

CPD practices and policies, CPD developed the Strategic Subject List (SSL), a computerized assessment tool that incorporates numerous sources of information to analyze crime as well as identifies and ranks individuals at risk of becoming a victim or possible offender in a shooting or homicide.⁶³ The tool was developed by the Illinois Institute of Technology and funded through the Department of Justice Bureau of Justice Assistance grant program, and some version of the tool has been used since 2012. When information on the SSL was first made public following the Freedom of Information Act (FOIA) litigation in 2017, the dataset included 398,684 individuals.⁶⁴

The SSL ranks and assigns risk tiers ranging from very low to very high⁶⁵ to individuals based on the following variables: the number of times an individual was a victim of a shooting; the individual's age during latest arrest; the number of times the individual was a victim of aggravated battery or assault; trends in criminal activity⁶⁶; the number of prior arrests for unlawful use of a weapon; the number of prior arrests for violent offenses; the number of prior narcotics arrests; and gang affiliation.⁶⁷ It is notable that a majority of these variables are based on arrest records, rather than convictions, which not only means that people who have not committed crimes may end up on the list but also that the list likely reflects CPD's unlawful and biased practices.⁶⁸ These facts were both confirmed by analysis

⁶³ See CHI. POLICE DEP'T, SUBJECT ASSESSMENT AND INFORMATION DASHBOARD (SAID), SPECIAL ORDER S09-11 (2019), <http://directives.chicagopolice.org/directives/data/a7a57b85-155e9f4b-50c15-5e9f-7742e3ac8b0ab2d3.html>; *Strategic Subject List*, CHI. DATA PORTAL, <https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List/4aki-r3np> (last visited Feb. 19, 2019).

⁶⁴ See *Strategic Subject List*, *supra* note 63; see also Brianna Posadas, *How Strategic Is Chicago's "Strategic Subject List"? Upturn Investigates.*, MEDIUM (June 22, 2017), <https://medium.com/equal-future/how-strategic-is-chicagos-strategic-subjects-list-upturn-investigates-9e5b4b235a7c>.

⁶⁵ When individuals are assessed as high risk, they are subject to heightened police scrutiny. See CHI. POLICE DEP'T, CUSTOM NOTIFICATIONS IN CHICAGO, SPECIAL ORDER S10-05 (2015), <http://directives.chicagopolice.org/directives/data/a7a57bf0-1456faf9-bfa14-570a-a2deebf33c56ae59.html>.

⁶⁶ This is not defined, and there is no publicly available information on how CPD calculates this number.

⁶⁷ This variable is included in the SSL dataset. And while the Mayor's office claims that this variable is no longer used, CPD confirmed its inclusion as a variable. See *Strategic Subject List - Dashboard*, CHI. DATA PORTAL, <https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List-Dashboard/wgnt-sjgb> (last visited Feb. 19, 2019); Yana Kunichoff & Patrick Sier, *The Contradictions of Chicago Police's Secretive List*, CHICAGO (Aug. 21, 2017, 8:44 AM), <https://www.chicagomag.com/city-life/August-2017/Chicago-Police-Strategic-Subject-List/>.

⁶⁸ SSL data shows that the arrests of people identified on the list overlap with areas that are heavily targeted by CPD for patrol, which is documented through the contact cards police fill out after an investigatory street stop. The areas that are subject to heightened CPD presence and SSL enforcement are concentrated in the South and West sides of Chicago, which are predominately non-white and heavily low-income neighborhoods. See *Strategic Subject List - Dashboard*, CHI.

of the SSL dataset. Independent analysis by Upturn and *The New York Times* found that more than one third of individuals on this list have never been arrested or a victim of a crime, and almost seventy percent of that cohort received a high risk score.⁶⁹ The SSL data also revealed that fifty-six percent of Black men under the age of thirty in Chicago have a risk score on the SSL, and this is the same demographic that has been disproportionately affected by CPD's unlawful and biased practices identified in the Department of Justice and ACLU reports.⁷⁰ These revelations are even more troubling in light of the conclusions of the only known validation study performed on an early version of the SSL by the RAND Corporation. The RAND study found the SSL was not successful in reducing gun violence or reducing the likelihood of victimization, inclusion on the SSL only had a direct effect on arrests, and the researchers noted these outcomes raised significant privacy and civil rights considerations.⁷¹

There are also concerns regarding how the SSL predictions and risk scores are used by CPD in the field. A CPD internal directive and public statements claim that the SSL is used to target individuals with social services as part of the custom notification procedure, which is part of a citywide violence interventional model.⁷² Yet the same CPD directive also encourages "[t]he highest possible charges" to be sought for any individuals on the SSL that received a custom notification and are subsequently

DATA PORTAL, <https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List-Dashboard/wgnt-sjgb> (last visited Feb. 19, 2019); *Map: Investigatory Stop Reports by District*, OFF. INSPECTOR GEN.: INFO. PORTAL (Aug. 20, 2018), <https://informationportal.igchicago.org/map-investigatory-stop-reports-by-beat-and-district/>; see also Heather Cherone, *Here's How Many Officers Are Patrolling Your Neighborhood*, DNAINFO: CHI. (Apr. 17, 2017, 3:48 PM), <https://www.dnainfo.com/chicago/20170417/logan-square/heres-how-many-officers-are-patrolling-your-neighborhood-watchdog/> (reporting that twelve percent of officers patrolling Chicago's police districts are assigned to the Harrison and Englewood districts on Chicago's West and South sides, an area home to about five percent of the city's population).

⁶⁹ See Posadas, *supra* note 64. These findings repudiate prior CPD official statements suggesting that the SSL is populated with individuals that were known for driving gun violence, and reporting by the *Chicago Sun-Times* indicating that the list includes individuals that have previously been arrested and fingerprinted in Chicago since 2013. See Mick Dumke & Frank Main, *A Look Inside the Watch List Chicago Police Fought to Keep Secret*, CHI. SUN-TIMES (May 18, 2017), <https://chicago.suntimes.com/politics/what-gets-people-on-watch-list-chicago-police-fought-to-keep-secret-watchdogs/>.

⁷⁰ See Kunichoff & Sier, *supra* note 67 (reporting on the SSL data); see also CIVIL RIGHTS DIV., U.S. DEP'T OF JUSTICE & U.S. ATTORNEY'S OFFICE FOR THE N. DIST. OF ILL., *supra* note 56, at 68, 143–50 (discussing CPD's targeting of young men of color); ACLU OF ILL., *supra* note 49, at 9.

⁷¹ See Jessica Saunders et al., *Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot*, 12 J. EXPERIMENTAL CRIMINOLOGY 347, 366–67 (2016).

⁷² See CHI. POLICE DEP'T, *supra* note 65; Dumke & Main, *supra* note 69 (“[Chicago officials] say they don’t rely on the scores alone when deciding which people to keep track of. Hundreds of people are flagged for interventions based on outstanding arrest warrants and ‘human intelligence’ in addition to their scores . . .”).

arrested,⁷³ and the RAND study observed that most police districts did not focus SSL enforcement on social service interventions.⁷⁴ CPD does not publicly release data on successful interventions, but available data and press coverage on CPD's SSL enforcement indicates arrests as a primary outcome, and in some cases a stated goal.⁷⁵ This disconnect from CPD's stated goals and outcomes for the SSL was also discussed in the RAND study, which suggested that the lack of a centralized crime prevention strategy and district-level guidance on how to use the SSL in the field may undermine any potential utility of the technology as a crime prevention strategy.⁷⁶ The study found that most CPD officers did not receive guidance on how to use the SSL predictions which resulted in officers merely increasing contacts with individuals on the list.⁷⁷ The RAND researchers noted that "it is not at all evident that contacting people at greater risk of being involved in violence—especially without further guidance on what to say to them or otherwise how to follow up—is this relevant strategy to reduce violence."⁷⁸ This observation was also supported by the study's results, which found increased police contacts with individuals on the SSL had no direct effect on arrest or victimization.⁷⁹

Despite these demonstrated concerns about the SSL and the underlying CPD practices and policies that still await reform, in 2017 CPD entered a contract with the University of Chicago Crime Lab to develop and

⁷³ See CHI. POLICE DEP'T, *supra* note 65.

⁷⁴ See Saunders et al., *supra* note 71, at 356 (noting that district level guidance on SSL enforcement only occurred in 18.7% of COMPSTAT meetings observed and that such guidance included "chang[ing] the focus from arresting SSL subjects for minor offenses (for which they would be immediately released) to finding ways to detain SSL subjects over the long term").

⁷⁵ See, e.g., Sam Charles, *30 Arrested in Raids Aimed at Curbing Memorial Day Weekend Violence*, CHI. SUN-TIMES (May 27, 2017, 12:57 PM), <https://chicago.suntimes.com/news/30-arrested-in-raids-aimed-at-curbing-memorial-day-weekend-violence/amp/> ("All those taken into custody so far, along with those still being sought, are on the department's [SSL] . . . 'Our goal was to identify the people that we think are driving the violence . . . and let them spend the weekend in Cook County Jail,' Riccio said."); Jeremy Gerner, *In Crackdown on Violence, Chicago Police Arrest More Than 100 in Gang Raids*, CHI. TRIB. (May 20, 2016, 7:39 PM), <https://www.chicagotribune.com/news/local/breaking/ct-chicago-police-crackdown-on-violence-met-20160520-story.html> ("Chicago police have carried out an extensive gang takedown, arresting more than 115 people on the department's 'strategic subject list'—those believed to be most prone to violence."); Kunichoff & Sier, *supra* note 67 (discussing how "in 2016, 1,024 notifications were attempted, 558 were completed, and only 26 people attended a call-in To put this in perspective, CPD has stated that 280 individuals with SSL scores were arrested in four gang raids during a six-month span in 2016").

⁷⁶ See Saunders et al., *supra* note 71, at 356, 367.

⁷⁷ *Id.* at 363 ("Individuals on the SSL were 50% more likely to have at least one contact card and 39% more likely to have any interaction (including arrest, contact cards, victimization, court appearances, etc.) with the Chicago PD than their matched comparisons in the year following the intervention.").

⁷⁸ *Id.* at 367.

⁷⁹ See *id.* at 363.

implement additional data-driven crime fighting strategies that will use predictive analytics and the SSL.⁸⁰ There is no evidence that this new initiative intends to account for ongoing consent decree negotiations or otherwise address CPD's unlawful and biased practices, including the dirty data generated by decades of these practices.

C. Case Study 2: New Orleans

The New Orleans Police Department (NOPD) has been investigated by the Department of Justice twice. The first investigation began in 1996 focusing on a wide range of police misconduct, but it ended without a consent decree in 2004 because NOPD pledged to reform itself.⁸¹ In 2010 at the invitation of then-Mayor Mitchell J. Landrieu,⁸² the Department reopened its investigation of NOPD, reviewing records.⁸³ The Department of Justice subsequently issued a report finding that NOPD engaged in a pattern or practice of excessive force disproportionately affecting Black residents; unlawful stops, searches, and arrests; failure to detect, prevent, or address bias-based profiling and other discriminatory policing on basis of race, national origin, and LGBT status; racial disparities in arrest rates and other police data; and gender discrimination in the failure to adequately respond to and investigate violence against women.⁸⁴ In 2013, the City of New Orleans and the Department of Justice entered a consent decree requiring structural and systemic reform with an independent monitor producing annual, quarterly, and special reports documenting compliance. These reports have mostly indicated good-faith, yet incremental progress, but the most recent Annual Report of the Consent Decree Monitor found NOPD in

⁸⁰ See UNIV. OF CHI. CRIME LAB, CHICAGO DEPARTMENT OF PROCUREMENT SERVICES NON-COMPETITIVE REVIEW BOARD (NCRM) APPLICATION (2017), <https://www.chicago.gov/content/dam/city/depts/dps/SoleSource/NCRB2017/ApprovedNCRBUOCcrimelab0217.pdf>.

⁸¹ Allen Johnson Jr., *What the Studies Said*, NEW ORLEANS MAG. (May 2011), <http://www.myneworleans.com/New-Orleans-Magazine/May-2011/WHAT-THE-STUDIES-SAID/> (discussing this history); Kimbriell Kelly et al., *Forced Reforms, Mixed Results*, WASH. POST (Nov. 13, 2015), https://www.washingtonpost.com/sf/investigative/2015/11/13/forced-reforms-mixed-results/?noredirect=on&utm_term=.e35d1149feb7 (same).

⁸² See Letter from Mitchell J. Landrieu, Mayor, City of New Orleans, to Eric H. Holder, Jr., Attorney Gen., U.S. Dep't of Justice (May 5, 2010), http://media.nola.com/crime_impact/other/LettertoAttyGenHolder.050510.pdf.

⁸³ Laura Maggi, *Federal Justice Department Officials Pledge Wide-Ranging Inquiry to Fix NOPD*, NOLA (May 18, 2010), https://www.nola.com/crime/2010/05/federal_justice_department_off.html (reporting that the Department of Justice opened an extensive investigation into NOPD's practices).

⁸⁴ See CIVIL RIGHTS DIV., U.S. DEP'T OF JUSTICE, INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT, at vi-xiii (2011), https://www.justice.gov/sites/default/files/crt/legacy/2011/03/17/nopd_report.pdf [hereinafter INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT].

non-compliance regarding stop, search, and arrest practices.⁸⁵

The extreme scope of NOPD's unlawful and biased practices between 2005 and 2011 is enough to cast doubt on all police and crime data, relevant to predictive policing systems, created during this period. In fact, the Department of Justice's Investigation Report, which substantiated the duration and scope of unlawful and biased practices, also documented evidence of "dirty data." The report identified several concerning disparities and inconsistencies in NOPD arrest and field interview card data (documenting NOPD encounters with residents, even those that do not result in arrest), and it questioned NOPD policies that encouraged unwarranted and potentially privacy-violating data collection.⁸⁶ For example, the report noted that in 2009, when NOPD arrest data was compared to national averages, "[t]he level of disparity for youth in New Orleans is so severe and so divergent from nationally reported data that it cannot plausibly be attributed entirely to the underlying rates at which these youth commit crimes, and unquestionably warrants a searching review and a meaningful response from the Department."⁸⁷ Additionally, the Department of Justice expressed concerns regarding omissions of essential information noting that NOPD "[p]olicies and practices for complaint intake do not ensure that complaints are complete and accurate, systematically exclude investigation of certain types of misconduct, and fail to track allegations of discriminatory policing."⁸⁸

In 2012, the City of New Orleans entered a pro bono contract with the data-mining firm Palantir to use its proprietary Gotham data analysis and profiling services for crime-forecasting and to inform public safety strategies deployed by NOPD and other public safety agencies.⁸⁹ There is limited

⁸⁵ OFFICE OF THE CONSENT DECREE MONITOR NEW ORLEANS, LA., 2017 ANNUAL REPORT OF THE CONSENT DECREE MONITOR FOR THE NEW ORLEANS POLICE DEPARTMENT CONSENT DECREE 9–11, 18 (2018), <http://nopdconsent.azurewebsites.net/Media/Default/Documents/Reports/550%20MONITORS%20ANNUAL%20REPORT%20OF%202017.pdf> (summarizing compliance-related findings after describing the scope of review performed over the past few years and noting noncompliance).

⁸⁶ See INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT, *supra* note 84, at 29–30.

⁸⁷ *Id.* at x.

⁸⁸ *Id.* at xviii.

⁸⁹ See Defendant Kentrell Hickerson's Motion for Leave to Supplement His *Second Motion for New Trial* with Additional Grounds at 1–2, *Louisiana v. Hickerson*, No. 516-272, 2018 WL 2009261, at *1–2 (La. Dist. Ct. Mar. 8, 2018) [hereinafter Defendant Kentrell Hickerson's Motion for Leave], <https://www.documentcloud.org/documents/4411697-Hickerson-Appeal-Defendant-s-Motion-to.html> (discussing this contract); Matt Sledge, *Convicted Gang Leader Can Challenge NOPD's Use of Crime-Fighting Software, Judge Rules*, NEW ORLEANS ADVOCATE (Mar. 14, 2018, 2:09 PM), https://www.theadvocate.com/new_orleans/news/courts/article_3a68a838-27bb-11e8-8b07-178e270926d4.html (reporting on a challenge to New Orleans's use of the Palantir technology). Presentation slides of two New Orleans government officials indicate that the Palantir system was used by NOPD to perform social network analysis and identify gang involvement in murders and shootings. It was also used by the Mayor's Office to locate neighborhood cleanup

information about the Palantir system and its partnership with City of New Orleans because the contract and its subsequent extensions were entered without the knowledge of key government officials and the public.⁹⁰ In fact, the lack of transparency about the use of the Palantir system has been the subject of a *Brady* challenge, citing the nondisclosure of the system's analysis about the defendant and related gang activities.⁹¹ Yet presentation materials of two New Orleans government officials indicate that the Palantir system relied on NOPD data, including calls for service, electronic police reports, field information cards, and crime lab analysis, as well as data gleaned from the City's criminal⁹² and non-law enforcement data sources, and open data sources, such as the location of liquor stores.⁹³ There is no indication from available government and vendor documents that the NOPD data used to implement the system was scrubbed for errors and irregularities or otherwise amended in light of the dirty data identified in the Department of Justice report. In fact, evidence suggests an elevated risk that the Palantir system relied on some form of NOPD's dirty data because the system's analysis reflected similar racial disparities and other biases of NOPD's practices and policies. City government documents highlighted that the NOPD system identified victims and perpetrators of violent or gang crimes as "overwhelmingly young, African American, male, undereducated, and

sites, and the Fire Department to increase its presence around schools in violent areas. See SARAH SCHIRMER, OFFICE OF MAYOR MITCH LANDRIEU, DEPLOYING PALANTIR GOTHAM IN NEW ORLEANS 11, 15, <https://assets.documentcloud.org/documents/4344815/Nola-hc3-Final-20140403.pdf>; see also PALANTIR TECHS., NOLA MURDER REDUCTION: TECHNOLOGY TO POWER DATA-DRIVEN PUBLIC HEALTH STRATEGIES 9, <https://www.documentcloud.org/documents/4344816-NOLA-Murder-Reduction-White-Paper.html> (describing the partnership between Palantir and New Orleans to identify "individuals exhibiting the highest predictors of violence").

⁹⁰ See Ali Winston, *Palantir Has Secretly Been Using New Orleans to Test Its Predictive Policing Technology*, VERGE (Feb. 27, 2018, 3:25 PM), <https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd>.

⁹¹ See Defendant Kentrell Hickerson's Motion for Leave, *supra* note 89 (bringing this challenge); Sledge, *supra* note 89 (reporting on the challenge's success); see also *Brady v. Maryland*, 373 U.S. 83, 87 (1963) (establishing the prosecutor's duty to turn over to the defense "evidence favorable to an accused" that is "material either to guilt or to punishment"). The hearings in Hickerson's case gave a rare glimpse into the NOPD predictive policing system. Relevant documents are on file with the author and in Louisiana state court at Orleans Parish Criminal Case No. 516-272.

⁹² Criminal data sources include, for example, calls for service, group and gang databases, Sheriff's Office arrest and booking records, and probation and parole data.

⁹³ The presentation slides of two New Orleans government officials indicate that the Palantir system was used to support criminal investigation, strategic homicide-reduction strategies, and to obtain indictments by employing multiple data sources. The presentation lists data sources including, but not limited to, jail calls and phone data, gang affiliations and violent activity, crime lab data, and social media. See SCHIRMER, *supra* note 89, at 14, 16; see also PALANTIR TECHS., *supra* note 89, at 9.

underemployed,” the same population that was disproportionately targeted by NOPD practices and overwhelmingly misrepresented in NOPD data.⁹⁴ Though there can be additional or alternative explanations for this correlation, the scope and severity of NOPD’s unlawful and biased practices and the extreme distortions identified in NOPD data suggests some level of attribution. The City of New Orleans has since cancelled its contract with Palantir in 2018, after public backlash regarding the secretive nature of the agreement.⁹⁵

D. Case Study 3: Maricopa County

In 2008, the ACLU filed a class-action lawsuit against the Maricopa County Sheriff’s Office (MCSO) for engaging in allegedly racially biased and unlawful police practices and policies as part of unlawful immigration enforcement operations,⁹⁶ which implicated contentious and unresolved legal questions regarding the authority and role of state and local police to enforce federal immigration laws.⁹⁷ The following year, the Department of Justice announced an investigation of MCSO, but the investigation was delayed because MCSO refused to provide access to pertinent material and personnel.⁹⁸ In 2011, the Department of Justice released an investigation findings letter documenting MCSO’s pattern of discriminatory behavior between 2007 and 2011, including discriminatory policing against Latino residents; unlawful stops and arrests; and unlawful retaliation against people

⁹⁴ SCHIRMER, *supra* note 89, at 6.

⁹⁵ See Ali Winston, *New Orleans Ends Its Palantir Predictive Policing Program*, VERGE (Mar. 15, 2018, 3:50 PM), <https://www.theverge.com/2018/3/15/17126174/new-orleans-palantir-predictive-policing-program-end>.

⁹⁶ First Amended Complaint, *Ortega Melendres v. Arpaio*, 598 F. Supp. 2d 1025 (D. Ariz. 2009) (No. CV 07-02513-PHX-MHN), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-first-amended-complaint>. Around January 2007, MCSO entered a Memorandum of Agreement with U.S. Immigrations and Customs Enforcement (ICE) that purports to authorize enforcement of federal immigration laws by specially nominated and cross-trained MCSO staff. ICE stipulated that the agreement did not authorize MCSO staff to perform any of the biased and unlawful practices alleged in the ACLU complaint, including random street operations targeting day laborers and using race or immigration status as pretext for unlawful traffic stops. *Id.* at 8–10.

⁹⁷ In its 2012 decision in *Arizona v. United States*, the Supreme Court held that states are preempted from arresting or detaining individuals on the basis of suspected removability under federal immigration law. 567 U.S. 387 (2012). However, this decision did not address the legality of immigration inquiries that arise during the normal course of police activities unrelated to immigration enforcement, or the limitations on data generated by these legally questionable police activities. The proliferation of Sanctuary City laws and President Trump’s aggressive and legally questionable immigration policies has further complicated these questions, and they currently remain unresolved.

⁹⁸ Letter from Thomas E. Perez, Assistant Attorney Gen., to Bill Montgomery, Maricopa Cty. Attorney 1 n.1 (Dec. 15, 2011), https://www.justice.gov/sites/default/files/crt/legacy/2011/12/15/mcso_findletter_12-15-11.pdf.

who make complaints or criticize the department.⁹⁹ The Department of Justice also noted concerns that MCSO practices created a “wall of distrust” that “substantially compromised effective policing by limiting the willingness of witnesses and victims to report crimes and speak to the police about criminal activity,” which affects crime data and public safety within the County.¹⁰⁰ In 2013, a federal court found that MCSO engaged in unconstitutional and racially biased traffic stops and detentions of Latino drivers, and it enjoined MCSO from enforcing policies that permitted unlawful immigration enforcement.¹⁰¹ In addition to detailing the scope of MCSO’s biased practices, the decision noted significant irregularities and omissions in MCSO’s records as well as evidence that MCSO officers and leadership openly, and often publicly, acknowledged biased and derogatory views and motives against Latino residents.¹⁰² In response, a federal judge issued a court order mandating an annual review of MCSO practices and data, in addition to requiring more specific reforms.¹⁰³

In 2015, the Department of Justice entered a settlement agreement¹⁰⁴ with MCSO addressing some of the unlawful practices identified in its 2011 investigation, and joined the ongoing ACLU lawsuit as a plaintiff. In 2016, a federal court found several MCSO officers in civil contempt for deliberately violating the 2013 court order and continuing to engage in unconstitutional and discriminatory practices.¹⁰⁵ In compliance with the 2013 federal court order, MCSO commissioned Arizona State University (ASU) to perform annual reviews of its data. Unsurprisingly, these reviews confirmed that this police data reflected the department’s unlawful and racially biased practices. The two existing ASU annual reports of MCSO data covering 2014 to 2017 revealed that even while under consent decree MCSO continued to engage in racially biased and unlawful traffic stops and arrests of Black and Latino residents, with Latinos experiencing a greater likelihood of post-stop arrests or searches from 2015 to 2016.¹⁰⁶

⁹⁹ *Id.* at 2.

¹⁰⁰ *Id.* at 16.

¹⁰¹ *Ortega Melendres v. Arpaio*, 989 F. Supp. 2d 822 (D. Ariz. 2013).

¹⁰² *Id.* at 830–31.

¹⁰³ *See Ortega Melendres v. Arpaio*, No. CV-07-02513-PHX-GMS, 2013 WL 5498218, at *7 (D. Ariz. Oct. 2, 2013), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-order>.

¹⁰⁴ Settlement Agreement, *United States v. Maricopa County*, No. 2:12-cv-00981-ROS (D. Ariz. July 17, 2015), <https://www.justice.gov/opa/file/631271/download>.

¹⁰⁵ *See Ortega Melendres v. Arpaio*, No. CV-07-02513-PHX-GMS, 2016 WL 2783715 (D. Ariz. May 13, 2016), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-2016-order>.

¹⁰⁶ DANIELLE WALLACE ET AL., ANNUAL REPORT FOR THE MARICOPA COUNTY SHERIFF’S OFFICE: YEARS 2016 TO 2017, at 8, 25–31 (2018), https://cvpcs.asu.edu/sites/default/files/content/projects/MCSOreport2016_2017_0.pdf; DANIELLE WALLACE ET AL., PRELIMINARY YEARLY REPORT FOR THE MARICOPA COUNTY

While it is clear that Maricopa County's police data reflects its history of biased policing practices, it is also a case where lack of transparency makes it difficult to know whether this data was specifically used in any local predictive policing system. There is no evidence of MCSO using its own predictive policing system, but four cities within Maricopa County, which share data and police resources with MCSO,¹⁰⁷ are actively using predictive policing software, or have previously participated in a predictive policing pilot that may have relied on MCSO data. In 2012, the Glendale Police Department in Maricopa County participated in a predictive policing pilot using RTMDx software, relying on Glendale police and crime data.¹⁰⁸ In this case, there is no evidence to suggest that MCSO data was directly used in this pilot. However, MCSO and the Glendale Police Department started officially sharing police and probation data in 2016 through AZ Link, a regional police data sharing platform which includes MCSO data dating from the period where the office is shown to have engaged in unlawful and biased practices and policies.¹⁰⁹ In 2016, the Mesa Police Department in Maricopa County entered a three-year contract with the predictive policing software company PredPol, which required the police department to provide local crime data.¹¹⁰ A 2011 Mesa City Council document reveals that the Mesa Police Department also uses AZ Link.¹¹¹ While it is evident that the Mesa Police Department uses MCSO data and that MCSO likely generates local police data as the backup law enforcement agency, it remains unconfirmed whether the department included this data directly in the crime

SHERIFF'S OFFICE, YEARS 2014–2015, at 28–31 (2016), <https://www.courthousenews.com/wp-content/uploads/2017/04/Maricopa-Bias-ASU-Report.pdf>.

¹⁰⁷ Many communities within Maricopa County rely on MCSO as its primary or backup law enforcement agency. The nature of this role is not articulated in available public documents, but Mesa, Tempe, parts of Peoria, and parts of Glendale are all listed as communities relying entirely or in part on MCSO services. See *Districts*, MARICOPA COUNTY SHERIFF'S OFF., <https://www.mcso.org/Patrol/Districts> (last visited Feb. 13, 2019).

¹⁰⁸ LESLIE KENNEDY, JOEL CAPLAN & ERIC PIZA, CONJUNCTIVE ANALYSIS REPORT: 2012 ROBBERY IN GLENDALE, AZ (2015), http://www.rutgerscps.org/uploads/2/7/3/7/27370595/gpd_conjanalysis.pdf.

¹⁰⁹ See Memorandum of Understanding between Maricopa County and City of Glendale (Sept. 27, 2016), <https://glendale-az.legistar.com/LegislationDetail.aspx?ID=2842290&GUID=9619C96A-1418-4455-A604-B85C399BE6AE&Options=&Search=>.

¹¹⁰ See CITY COUNCIL OF THE CITY OF MESA, ARIZ., CITY COUNCIL REPORT FOR AUGUST 22, 2016 (2016), <https://mesa.legistar.com/LegislationDetail.aspx?ID=2808201&GUID=46F97E21-FFAB-41E1-846A-4F6C4B60F6E0&Options=&Search=> (requesting approval of the contract); Maria Polletta, *Can New Mesa Police Tool Prevent Crime from Happening?*, AZCENTRAL (Oct. 21, 2016), <https://www.azcentral.com/story/news/local/mesa/2016/10/21/mesa-police-tool-prevent-crime-happening/89231252/>.

¹¹¹ See CITY COUNCIL OF THE CITY OF MESA, ARIZ., COUNCIL MINUTES FOR AUGUST 29, 2011 (2011), <http://apps.mesaaz.gov/meetingarchive/archiveddocuments/ClerkDetailedMinutes/August%2029,%202011%20RGL.pdf> (recommending the purchase of additional software for AZ Link).

data it provided to PredPol. So while there is no current evidence of dirty data in any of these systems, there is some risk that any system trained via AZ Link may be tainted by such data, depending on which variables it uses. And again, there is no evidence to suggest that either police departments or predictive vendors are prepared to address this concern.

In 2014, the Tempe Police Department, in Maricopa County, received a Department of Justice Bureau of Justice Assistance grant to create a person- and place-based predictive policing pilot using several vendors' software and services.¹¹² Though there is a study assessing the efficacy of this pilot, it does not indicate when the pilot occurred or the data the system used.¹¹³ However, as the Chicago case study shows, person-based predictive systems run an extremely high risk of having dirty data influence predictions.

Finally, the Peoria Police Department in Maricopa County has used HunchLab predictive policing software since 2015, but there are no publicly available documents detailing the data the software is using or the City's data sharing policies with MCSO.¹¹⁴

In sum, it is difficult to make definitive conclusions regarding the use of MCSO's dirty data in these predictive policing systems because of the lack of publicly available information on the implementation of these systems in each jurisdiction and uncertainty regarding the role of and relevant policies governing MCSO as a primary or backup law enforcement agency in each of these cities. Yet this case study does highlight important concerns about the extraterritorial nature of police data, particularly when the practices and policies of relevant police departments are ill-defined and implicate controversial legal questions regarding the authority of local police.

II

THE PROBLEMS OF USING POLICE DATA AS CURRENTLY CONSTITUTED

A. *Confirmation Feedback Loop*

As the above examples show, numerous jurisdictions suffer under ongoing and pervasive police practices replete with unlawful, unethical, and biased conduct. This conduct does not just influence the data used to build

¹¹² See BUREAU OF JUSTICE ASSISTANCE, U.S. DEP'T OF JUSTICE, CASE STUDY: TEMPE, ARIZ., POLICE DEPARTMENT (2017), <https://it.ojp.gov/CAT/Documents/CaseStudyTempeArizonaPoliceDepartment.pdf>.

¹¹³ *Id.* at 6.

¹¹⁴ See CITY OF PEORIA, ARIZ., FY16 COMPREHENSIVE ANNUAL FINANCIAL REPORT (2016); see also Andrew Bernier, *How Peoria Police Are Using New Technology to Better Predict When and Where Crime Will Happen*, KJZZ (Feb. 24, 2016), <https://science.kjzz.org/content/269439/how-peoria-police-are-using-new-technology-better-predict-when-and-where-crime-will>.

and maintain predictive systems; it supports a wider culture of suspect police practices and ongoing data manipulation. Add to this the lack of oversight and accountability measures regarding police data collection, analysis, and use, and it becomes clear that any predictive policing system trained on or actively using data from jurisdictions with proven problematic conduct cannot be relied on to produce valid results without extensive independent auditing or other accountability measures. Yet police technology vendors have shown no evidence of providing this accountability and oversight, and other governmental actors rarely have the tools to do so. Thus, in such jurisdictions, these systems should be met with considerable suspicion that they are neutral, unbiased, or without risks of discrimination, and they should not replace or otherwise circumvent police reform measures.¹¹⁵

Though there is research that empirically demonstrates that the mathematical models of predictive policing systems are susceptible to runaway feedback loops, where police are repeatedly sent back to the same neighborhoods regardless of the actual crime rate, such feedback loops are also a byproduct of the biased police data.¹¹⁶ More specifically, police data can be biased in two distinct ways. First and fundamentally, police data reflects police practices and policies.¹¹⁷ If a group or geographic area is disproportionately targeted for unjustified police contacts and actions, this group or area will be overrepresented in the data, in ways that often suggest greater criminality. Second, the data may omit essential information as a result of police practices and policies that overlook certain types of crimes and certain types of criminals.¹¹⁸ For instance, police departments, and predictive policing systems, have traditionally focused on violent, street, property, and quality of life crimes.¹¹⁹ Meanwhile, white collar crimes are comparatively under-investigated and overlooked in crime reporting, despite a strong probability that they occur at a greater frequency than some of the other crime categories combined. Research on white collar crime is limited because the scope and nature of the crimes are constantly evolving so studies and surveys are often too narrow in focus, and complaints fail to reach or are not investigated by law enforcement. However, available studies estimate

¹¹⁵ See, e.g., DILLON REISMAN, JASON SCHULTZ, KATE CRAWFORD & MEREDITH WHITTAKER, AI NOW INST., ALGORITHMIC IMPACT ASSESSMENTS: A PRACTICAL FRAMEWORK FOR PUBLIC AGENCY ACCOUNTABILITY (2018), <https://ainowinstitute.org/aiareport2018.pdf>.

¹¹⁶ See Danielle Ensign et al., *Runaway Feedback Loops in Predictive Policing*, 81 PROC. MACHINE LEARNING RES. 1 (2018), <https://arxiv.org/pdf/1706.09847.pdf>.

¹¹⁷ See Armacost, *supra* note 23, at 474 (discussing how this influences police data).

¹¹⁸ See Tom Meagher, *13 Important Questions About Criminal Justice We Can't Answer*, MARSHALL PROJECT (May 15, 2016, 10:00 PM), <https://www.themarshallproject.org/2016/05/15/13-important-questions-about-criminal-justice-we-can-t-answer?ref=hp-1-112#.KBW3hNgLs> (listing a number of these omissions).

¹¹⁹ See, e.g., HUNCHLAB, *supra* note 38, at 3 “[M]ost clients use HunchLab to model major crimes like burglaries, robberies, assaults, homicides, and motor vehicle thefts”).

that approximately 49% of businesses and 25% of households have been victims of white collar crimes, compared to a 1.06% prevalence rate for violent crimes and a 7.37% prevalence rate for property crime.¹²⁰ Thus, while there is a significant need for more research on white collar crimes, available data demonstrates that these crimes occur at a greater frequency than crimes traditionally targeted by police departments and prominent predictive policing, like property and violent crimes.

The confluence of these distinct forms of skewed inputs ends up producing a questionable data-driven justification for increased policing and surveillance of historically overpoliced communities, and in turn reinforces popular misconceptions regarding the criminality and safety of underrepresented individuals and communities.¹²¹ The impact of these problems is most salient in the case study of New Orleans. The report on the Department of Justice investigation of the New Orleans Police Department revealed that officers improperly targeted and arrested transgender residents, sometimes fabricating evidence of a crime as well as exploiting archaic and biased statutes like “crimes against nature [by solicitation]”—a statute that criminalizes sexual conduct that is considered morally unacceptable and requires registration as a sex offender.¹²² The Department of Justice found that in addition to these practices being discriminatory, they also raised significant concerns about the “efficient and effective use of resources to ensure public safety” since individuals convicted of “crimes against nature” made up approximately 40% of the jurisdiction’s sex offender registry and the police department was charged with monitoring all registrants’ compliance.¹²³ Moreover, since 80% of those registrants were also Black, the Department of Justice suggested there was an element of racial bias as well, which was confirmed by community members who told investigators “they

¹²⁰ See RACHEL E. MORGAN & GRACE KENA, BUREAU OF JUSTICE STATISTICS, U.S. DEP’T OF JUSTICE, NCJ 252121, CRIMINAL VICTIMIZATION, 2016: REVISED 11 (2018), <https://www.bjs.gov/content/pub/pdf/cv16re.pdf> (reporting these figures); RODNEY HUFF, CHRISTIAN DESILETS & JOHN KANE, NAT’L WHITE COLLAR CRIME CTR., THE 2010 NATIONAL PUBLIC SURVEY ON WHITE COLLAR CRIME 14 (2010), https://www.nw3c.org/docs/research/2010-national-public-survey-on-white-collar-crime.pdf?sfvrsn=e51bbb5d_8 (same); PRICEWATERHOUSECOOPERS, PULLING FRAUD OUT OF THE SHADOWS 5 (2018), <https://www.pwc.com/gx/en/services/advisory/forensics/economic-crime-survey.html> (same); Gerald Cliff & April Wall-Parker, *Statistical Analysis of White-Collar Crime*, OXFORD RES. ENCYCLOPEDIA CRIMINOLOGY (Apr. 2017), <http://criminology.oxfordre.com/view/10.1093/acrefore/9780190264079.001.0001/acrefore-9780190264079-e-267> (same).

¹²¹ See, e.g., Cynthia Lum, *The Influence of Places on Police Decision Pathways: From Call for Service to Arrest*, 28 JUST. Q. 631, 632 (2011) (“Place-based cues, especially those most noticeable to an officer (e.g., socioeconomic status, poverty, racial and ethnic makeup, disorder, crime, pedestrian and traffic density, and land use), may significantly affect an officer’s worldview and thereby his or her discretion.”).

¹²² INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT, *supra* note 84, at x.

¹²³ *Id.*

believe some officers equate being African American and transgendered [sic] with being a prostitute.”¹²⁴ The report also recognized the long-term consequences of the police department’s unlawful and biased practices, noting, “for the already vulnerable transgender community, inclusion on the sex offender registry further stigmatizes and marginalizes them, complicating efforts to secure jobs, housing, and obtain services at places like publicly-run emergency shelters.”¹²⁵

Confirmation feedback loops are so pernicious because they obfuscate the realities of crime and public safety. This is often magnified by public perceptions and public policy. When people observe increased police presence or contacts in marginalized communities it can reinforce unwarranted assumptions and stereotypes.¹²⁶ Indeed, continued exposure to or reinforcement of these stereotypes, especially in the absence of a counternarrative, can allow society to maintain a prejudice against marginalized groups while still maintaining an explicit commitment to egalitarianism.¹²⁷ These complex yet contradictory sentiments can incite responses that perpetuate this feedback loop. For instance, observing racially biased police practices can reinforce racial animus and false stereotypes of violence and criminality of certain racial or ethnic groups, which can result in improper calls for service for non-criminal activity that is perceived as suspicious or causes discomfort.¹²⁸ This is well documented in research on

¹²⁴ *Id.*; see also Leonore F. Carpenter & R. Barrett Marshall, *Walking While Trans: Profiling of Transgender Women by Law Enforcement, and the Problem of Proof*, 24 WM & MARY J. RACE, GENDER & SOC. JUST. 5 (2017) (describing the harms caused by police regularly profiling transgender women, particularly those of color, as sex workers and the difficulty they face in proving this biased police practice); Michael J. Griffin, Note, *Intersecting Intersectionalities and the Failure of the Law to Protect Transgender Women of Color in the United States*, 25 TUL. J.L. & SEXUALITY 123, 135 (2016) (“Police will also often assume that transgender people in certain areas are engaging in sex work, and will use dubious evidence of alleged crimes, or the intent to commit them, to arrest transgender people . . .”).

¹²⁵ INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT, *supra* note 84, at x.

¹²⁶ See ANGELA J. HATTERY & EARL SMITH, *POLICING BLACK BODIES: HOW BLACK LIVES ARE SURVEILLED AND HOW TO WORK FOR CHANGE* 215–17, 232–34 (2018) (discussing how stereotyping of Blacks leads to racial profiling and wrongful convictions).

¹²⁷ See JILLIAN OLINGER, KELLY CAPATOSTO & MARY ANA MCKAY, *CHALLENGING RACE AS RISK* 21 (2017), <http://kirwaninstitute.osu.edu/wp-content/uploads/2017/02/implicit-bias-housing.pdf> (describing the causes and effects of implicit bias).

¹²⁸ See HATTERY & SMITH, *supra* note 126, at 215; Elijah Anderson, “*The White Space*,” 1 SOC. RACE & ETHNICITY 10, 13 (2015) (“In the absence of routine social contact between blacks and whites, stereotypes can rule perceptions, creating a situation that estranges blacks. . . . In other words, whites and others often stigmatize anonymous black persons by associating them with the putative danger, crime, and poverty . . .”); Elise C. Boddie, *Racial Territoriality*, 58 UCLA L. REV. 401, 434–46 (2010) (discussing how racial territoriality is a product of conscious and implicit biases that instigate territorial behavior that seeks to exclude people of color based on how they are perceived or represented); Angela Onwuachi-Willig, *Policing the Boundaries of Whiteness: The Tragedy of Being “Out of Place” from Emmett Till to Trayvon Martin*, 102 IOWA L. REV. 1113, 1170–83 (2017) (describing how white residents of the neighborhood where Trayvon Martin was

the social phenomena of “shopping while Black”¹²⁹ in addition to the recent onslaught of media reports of white residents calling the police on Black residents for non-criminal and innocuous actions like barbecuing in a park or not smiling at a white neighbor.¹³⁰ This societal response of internalized bias and feedback loops is also becoming more prevalent as historically segregated and majority non-white and lower income neighborhoods experience gentrification. There is a growing body of evidence documenting heightened neighbor-driven police enforcement in gentrifying neighborhoods.¹³¹ Most recently, a study on 311 calls for service data in New York City found that lower-income communities of color with the largest influxes of white residents experienced significantly higher increases in quality of life complaints, and summons and arrests outcomes were three times more likely than in neighborhoods without large influxes of white residents.¹³²

These societal responses to the feedback loop contribute to policing’s dirty data problem in several distinct ways, but the most concerning influence, especially in the predictive policing context, is that calls for

murdered perceived Black people as intruders which resulted in repeated 911 calls despite the lack of evidence to support such suspicions).

¹²⁹ See, e.g., Shaun L. Gabbidon, *Racial Profiling by Store Clerks and Security Personnel in Retail Establishments*, 19 J. CONTEMP. CRIM. JUST. 345 (2003); Cassi Pittman, “Shopping While Black”: Black Consumers’ Management of Racial Stigma and Racial Profiling in Retail Settings, J. CONSUMER CULTURE (2017); George E. Schreer, Sandra Smith & Kirsten Thomas, “Shopping While Black”: Examining Racial Discrimination in a Retail Setting, 39 J. APPLIED SOC. PSYCHOL. 1432 (2009).

¹³⁰ See, e.g., Carla Herreria, *Woman Calls Police on Black Family for BBQing at a Lake in Oakland*, HUFFPOST (May 18, 2018), https://www.huffpost.com/entry/woman-calls-police-oakland-barbecue_us_5af50125e4b00d7e4c18f741 (describing how police detained and questioned a Black family after a white woman called the police regarding the family’s attempt to barbeque at a public lake in Oakland, California); Amber Jamieson, *A White Woman Called the Police on Bob Marley’s Granddaughter for Not Smiling at Her*, BUZZFEED NEWS (May 10, 2018), <https://www.buzzfeednews.com/article/amberjamieson/black-artists-airbnb-white-woman-police-cops#rwYRXPYOg> (describing how police detained three Black people after a white resident called the police because the individuals failed to smile, wave back, and acknowledge her); P.R. Lockhart, *A Black Political Candidate Was Canvassing in Her District. Then Someone Called the Police.*, VOX (Sept. 20, 2018), <https://www.vox.com/identities/2018/9/20/17883018/shelia-stubbs-canvassing-911-police-racial-profiling-wisconsin> (describing how police received a false complaint of suspected drug activity by a Black legislator canvassing the area).

¹³¹ See, e.g., Ayobami Laniyonu, *Coffee Shops and Street Stops: Policing Practices in Gentrifying Neighborhoods*, 54 URB. AFF. REV. 898 (2018) (presenting empirical analysis demonstrating a strong positive association between gentrification and the adoption of punitive policing strategies); Lam Thuy Vo, *They Played Dominoes Outside Their Apartment for Decades. Then the White People Moved In and Police Started Showing Up.*, BUZZFEED NEWS (June 29, 2018), <https://www.buzzfeednews.com/article/lamvo/gentrification-complaints-311-new-york#xoJoxzbEA>; ANTI-EVICTION MAPPING PROJECT, 311 REPORTS IN SF BY NEIGHBORHOOD 2008–2016, <http://www.antievictionmappingproject.net/311.html> (noting a disproportionate increase in 311 calls in gentrifying neighborhoods in San Francisco).

¹³² *New Neighbors and the Over-Policing of Communities of Color*, COMMUNITY SERVS. SOC’Y N.Y. (Jan. 6, 2019), <http://www.cssny.org/news/entry/New-Neighbors>.

service provide several opportunities for discretion and selective enforcement, which can further distort police data. There is not only a great amount of subjectivity permitted in assessing the validity of a call for service and whether to involve law enforcement, but once police are present, they have full discretion in how to respond (e.g., whether to arrest individuals involved in the incident or de-escalate) and how to report their interactions.¹³³ This includes whether to classify their interactions as crimes, infractions, or non-criminal incidents. All of these subjective decisions alter police data that is often used in predictive policing systems (e.g., crime and calls for service data), and both longstanding research on police officer discretion and more recent research on calls for service outcomes demonstrate that police bias and institutional interests are reflected in the outcomes and reporting, despite actual crime levels and neighborhood conditions.¹³⁴ Thus, when the dirty data generated by dubious calls for services and subjective police enforcement and reporting is used in predictive policing systems, the technology can produce predictions that further perpetuate confirmation feedback loops.

Confirmation feedback loops also influence public policy by driving or providing justification for government officials to support policies that attempt to micromanage or push out communities that are misperceived as producing problems or increasing disorder. This phenomenon of artificially manufactured moral panic and public consent to new forms of state control was interrogated in the groundbreaking book, *Policing the Crisis*.¹³⁵ Stuart Hall and his co-authors conducted an empirical study of the social construction of street crime, and the societal labeling of Black men as “muggers.” The authors found that despite public belief that there was a new street crime pandemic, there was little evidence to support this belief.

¹³³ See Lum, *supra* note 121, at 640, 643 (illustrating police officer decision pathways in the context of calls for service).

¹³⁴ See, e.g., ERIC J. SCOTT, NAT’L INST. OF JUSTICE, U.S. DEP’T OF JUSTICE, CALLS FOR SERVICE: CITIZEN DEMAND AND INITIAL POLICE RESPONSE (1981), <https://www.ncjrs.gov/pdffiles1/Digitization/78362NCJRS.pdf> (analyzing operators’ responses to calls for service); David A. Klinger & George S. Bridges, *Measurement Error in Calls-for-Service as an Indicator of Crime*, 35 CRIMINOLOGY 705, 721–24 (1997) (concluding that estimates of neighborhood crime rates based on calls for service and police dispatch data are biased and can be misleading because the data is embedded with errors and subject to undercount based on neighborhood or community characteristics); Lum, *supra* note 121, at 635 (“Places with high proportions of Black residents may be treated differently by officers than the places that are predominately White, even if both places have similar levels of social disorganization or crime.”); Douglas A. Smith, Christy A. Vishner & Laura A. Davidson, *Equity and Discretionary Justice: The Influence of Race on Police Arrest Decisions*, 75 J. CRIM. L. & CRIMINOLOGY 234 (1984) (finding that in police encounters in which officers exercise greater discretion, Black people were more likely to be arrested than white people); Vo, *supra* note 131 (finding that police responses to 311 calls increased sixfold as a Harlem neighborhood experienced increased gentrification).

¹³⁵ See generally STUART HALL ET AL., *POLICING THE CRISIS: MUGGING, THE STATE, AND LAW AND ORDER* (1978).

Instead, they argued, street crime was not new but manufactured as a new problem by media, which then influenced which people police identified as criminals, and reinforced biases of judges who created justifications for the state-sponsored control of the Black community.

Now this phenomenon is commonly experienced through austerity measures and policies that criminalize the conditions that contribute to the marginalization of some communities over others.¹³⁶ Common examples are nuisance laws and ordinances, which empower municipal governments to penalize individuals and communities for a certain number of calls for service or alleged “nuisance” conduct, which is an ill-defined category of conduct that can range from assault to littering depending on the jurisdiction.¹³⁷ A recent New York Civil Liberties Union report found that these policies disproportionately affect poor communities of color because they “amplify the harms of the criminal justice system and exacerbate socioeconomic and racial inequalities by making housing instability a consequence of law enforcement.”¹³⁸ These policies and practices result in the displacement of these communities and are often surreptitiously pursued in order to attract private investment and consumption.¹³⁹

B. *The Boundless Risks of Biased Police Data*

These case studies demonstrate that without an empowered and independent authority, the potentially unlawful and biased practices and policies of police departments as well as the subsequent data produced through these practices can remain unaddressed and uncorrected. If dirty data is fed into a new predictive system, it can fundamentally taint its

¹³⁶ See Casey Kellogg, *There Goes the Neighborhood: Exposing the Relationship Between Gentrification and Incarceration*, 3 THEMIS: RES. J. JUST. STUD. & FORENSIC SCI. 177, 184 (2015) (“As migration patterns lead to more homogenous communities, it becomes easier to label the behavior of racial and class groups as criminal.”).

¹³⁷ See N.Y. CIVIL LIBERTIES UNION, MORE THAN A NUISANCE: THE OUTSIZED CONSEQUENCES OF NEW YORK’S NUISANCE ORDINANCES 6 (2018), https://www.nyclu.org/sites/default/files/field_documents/nyclu_nuisancereport_20180809.pdf (describing the range of conduct that can be included in a nuisance ordinance); Anna Kastner, *The Other War at Home: Chronic Nuisance Laws and the Revictimization of Survivors of Domestic Violence*, 103 CALIF. L. REV. 1047, 1052 (2015) (“Nuisance law is a well-established, but vague, common law doctrine. According to Prosser, ‘[t]here is general agreement that it is incapable of any exact or comprehensive definition,’ because it contains a mixture of criminal law, tort law, and property law.”).

¹³⁸ N.Y. CIVIL LIBERTIES UNION, *supra* note 137, at 10.

¹³⁹ See Donald C. Bryant, Jr. & Henry W. McGee, Jr., *Gentrification and the Law: Combatting Urban Displacement*, 25 WASH. U. J. URB. & CONTEMP. L. 43, 70–72 (1983) (discussing how housing codes can be used to increase compliance costs, pushing out tenants unable to afford rent increases); Jefferson, *supra* note 22, at 3–4 (2018) (“[C]ity governances vying to become magnets of consumption, finance, gentrification, and investment have fashioned policing policies designed to micromanage the public presence of problem populations and the disorder . . . that ails them.”); Kellogg, *supra* note 136, at 192–95.

recommendations. This can further ingrain biases in supposedly “neutral” systems. This is important since there are few political and institutional incentives that encourage self-monitoring and reform, or ongoing auditing of data systems.¹⁴⁰ We can see this in the failure of NOPD to self-reform after the first Department of Justice investigation into the department’s practices, and similarly in MCSO’s persistent defiance of the federal court order.

These case studies also demonstrate that merely identifying unlawful and biased practices is not enough. The data collection, analysis, and use of these practices must be reformed as well. Though most of the jurisdictions reviewed in our research engaged in some level of data collection and review reforms, none of the legal agreements restricted the use of the data generated during the periods of unlawful and biased police practices, which would be a meaningful limitation in future legal agreements on police reform. Thus, restrictions or prohibitions on the use of the historical data generated by unlawful and biased practices are necessary to ensure that the legacy of such practices is not perpetuated through the systems that rely on such data.¹⁴¹

Moreover, police data generated by the unlawful or biased practices and policies of a specific police department or division can corrupt practices and data in other jurisdictions, and skew decision-making throughout the criminal justice system, often in ways that are difficult to account for and correct. The Maricopa County case study demonstrates additional risks from dirty data production—the risks that other jurisdictions, including ones that have not been found to systematically violate the law, may incorporate such data into their own predictive policing systems, potentially corrupting additional new data and practices. Data sharing between police departments occurs frequently and in some cases is encouraged and given federal government support.¹⁴² And data sharing is not limited to law enforcement

¹⁴⁰ See, e.g., Friedman & Ponomarenko, *supra* note 25, at 1831 (highlighting why police are immune from regulation or constraints by other branches of government and other unique challenges in regulating police); Stephen Rushin, *Using Data to Reduce Police Violence*, 57 B.C. L. REV. 117, 152–54 (2016) (describing how municipal unwillingness to allocate resources for police reform, collective bargaining, and civil service protections inadvertently discourage police leadership from proactively and forcefully responding to misconduct); Myriam E. Gilles, *Breaking the Code of Silence: Rediscovering Custom in Section 1983 Municipal Liability*, 80 B.U. L. REV. 17, 31 (2000) (“Aside from the overdeterrence of individual officers, it seems clear that where liability falls solely on individual officers, municipalities have little incentive to develop comprehensive responses to rampant unconstitutional practices.”).

¹⁴¹ See, e.g., Mark Puente, *LAPD Data Programs Need Better Oversight to Protect Public, Inspector General Concludes*, L.A. TIMES (Mar. 12, 2019, 5:00 AM), <https://www.latimes.com/local/lanow/la-me-ln-lapd-data-20190312-story.html> (summarizing a report that called for greater oversight over LAPD’s data-driven computer programs in order to combat unfair arrests and detentions).

¹⁴² See RACHEL LEVINSON-WALDMAN, BRENNAN CTR. FOR JUSTICE, WHAT THE GOVERNMENT DOES WITH AMERICANS’ DATA 3 (2013), <https://www.brennancenter.org/publication/what-government-does-americans-data> (noting that federal law and agency directives often encourage the sharing of data between local and federal

agencies. Police and crime data are used in decisionmaking during prosecution, pretrial services, adjudication, sentencing, and corrections, as well as non-criminal justice related political decisions, such as community investment and development.¹⁴³ The lack of transparency and oversight regarding police practices, policies, and the data created through these practices and policies raises serious concerns regarding the possibility that any predictive system relying on police data could operate in a fair and just manner.

CONCLUSION

Data is seen as an important tool for policymaking and governance because in its absence there is often too much reliance on subjective factors. The last twenty years have seen a widespread adoption of data-driven practices, policies, and technologies in the public sector.¹⁴⁴ Yet this increasing reliance on data to assess and make decisions about complicated social, economic, and political issues presents serious risks to fairness, equity, and justice, if greater scrutiny is not given to the practices underlying the creation, auditing, and maintenance of data.

Our research demonstrates the risks and consequences associated with overreliance on unaccountable and potentially biased data to address sensitive issues like public safety. These case studies show that illegal police practices can significantly distort the data that is collected, and the risks that dirty data will still be used for law enforcement and other purposes. The failure to adequately interrogate and reform police data creation and collection practices elevates the risks of skewing predictive policing systems and creating lasting consequences that will permeate throughout the criminal justice system and society more widely.

There may be a natural inclination to assume that predictive policing vendors can address the problems of dirty data identified in this study by removing known cases. But such mitigation methods are likely inadequate for a number of reasons. First, if there are few incentives and almost no requirements for police departments to self-monitor and reform practices or policies that create biased or dirty data, it is unlikely that police departments would identify these problems for a vendor to remove or otherwise address. Second, there is no current methodology or mechanism for identifying these problematic practices and policies in real-time; therefore, any system that

authorities, especially with regard to terrorism); see, e.g., *About the RISS Program*, REGIONAL INFO. SHARING SYS., <https://www.riss.net/about-us/> (last visited Feb. 14, 2019) (describing the mission of the federally sponsored Regional Information Sharing Systems Program as “to assist local, state, federal, and tribal criminal justice partners by providing adaptive solutions and services that facilitate information sharing”).

¹⁴³ See Jefferson, *supra* note 22.

¹⁴⁴ See LEVINSON-WALDMAN, *supra* note 142 (discussing this history).

includes recent or live data may be subject to additional undocumented biases. Third, as we have argued, a fundamental flaw of police data is that it does not capture all relevant crime information because of institutional policies or practices that ignore certain types of crimes or criminals, negative community relations that affect which crimes the police track, and corrupt or unethical practices that lead to the omission or manipulation of police records. There is no documented practice demonstrating meaningful ways for a vendor to adjust its system for what is unknown or not recorded. The absence of data is as significant as its creation, yet there is no technical “fix” for this. Instead, mitigation efforts should be focused on developing reliable mechanisms for assessing the harms inherent in the use of historical police data, as well as data generated after implementation of police data collection reforms, and backed by strong public transparency and accountability measures.

The jurisdictions researched for this paper were limited to police departments that were subjects of publicized investigations and federal litigation. These case studies demonstrate the importance of independent government investigations and federal court litigation in uncovering unlawful and biased police practices that would otherwise persist without federal government intervention. Yet these crucial mechanisms for uncovering and addressing problematic police practices have been threatened with the parting acts of former U.S. Attorney General Jeff Sessions just before his unexpected forced resignation.¹⁴⁵ Before he left office, he issued a Department of Justice policy memo significantly limiting the use of consent decrees by requiring top political appointees to sign off, limiting their scope and duration, and requiring department attorneys to provide evidence of additional violations *beyond* unconstitutional behavior.¹⁴⁶ These limitations are significant and serious. The result may be that problematic police departments will remain unchecked, not because of lack of evidence of unconstitutional practices, but because the new standard of evidence is extremely high or because political leadership refuses to sign off. In light of these developments and the absence of incentives for self-scrutiny and reform, collective action for greater accountability, oversight, and redress is urgent. A broad coalition of stakeholders is needed to push public discourse on the drivers and consequences of dirty data, and to motivate government officials to act to ensure that principles of fairness, equity, and justice are reflected in government practices.

¹⁴⁵ See Katie Benner, *Sessions, in Last-Minute Act, Sharply Limits Use of Consent Decrees to Curb Police Abuses*, N.Y. TIMES (Nov. 8, 2018), <https://www.nytimes.com/2018/11/08/us/politics/sessions-limits-consent-decrees.html>.

¹⁴⁶ *Id.*

APPENDIX

For this review, we examined available government and legal documents to determine the scope and severity of police practices, policies, and data issues, as well as correlated time periods and geographic scope. These documents included government investigation reports, supplemental investigation documents, court filings, court orders, consent decrees, and post-settlement commissioned reports. We examined available public documents about predictive policing systems—marketing materials, white papers, peer-reviewed studies, and vendor blogs and websites—to understand the methodologies and epistemologies informing the development of the systems, types of data used by the systems, and anticipated outcomes. Finally, we reviewed available public data about the use or proposed use of predictive policing systems within the selected jurisdictions, including media reports, local government records, relevant government policies, and studies or reviews on the use of the system within a given jurisdiction.

| Jurisdiction | Type of Agreement (Year) | Unlawful/Biased Practices | Timeframe of Unlawful/Biased Practices | Predictive Policing System (Year) | Evidence of Biased Data in Predictive Policing |
|----------------|--|--|--|--|---|
| Baltimore (MD) | DOJ Consent Decree (2017) ¹⁴⁷ | Unconstitutional and racially biased stops, searches, and arrests; excessive use of force; retaliation against individuals for constitutionally-protected expression | 2010–2016 | Baltimore City (in development); Baltimore County (several pilots) | Unclear. Evidence suggests BPD and County Police shared data that could have been used in the County's pilots. BPD plans to acquire or develop a predictive |

¹⁴⁷ Consent Decree, *supra* note 10.

| | | | | | |
|---------------|---|---|-----------|---|--|
| | | | | | policing system. |
| Boston (MA) | Boston Police Department (BPD) and ACLU Commissioned Investigation (2015) ¹⁴⁸ | Racially discriminatory stop and frisk practices | 2007–2010 | NIJ-grant (2009 retrospective review of Safe Streets Teams); 2010 BPD official stated plans to develop a system | Safe Streets Teams system used 2000–2009 BPD data. |
| Chicago (IL) | Settlement Agreement (2015), ¹⁴⁹ DOJ Memorandum of Agreement (2018), ¹⁵⁰ IL Attorney General Consent Decree (2019) ¹⁵¹ | Pattern and practice of unconstitutional and racially biased use of force; poor data collection to identify and address unlawful conduct | 2011–2016 | RTMDx (2012); Strategic Subjects Lists | Both systems used arrest and other police data. |
| Ferguson (MO) | DOJ Consent Decree (2016) ¹⁵² | Pattern or practice of unlawful and discriminatory stops, searches, and arrests; excessive use of force; First Amendment violations; DOJ noted similar problems in St. Louis County | 2007–2016 | HunchLab (St. Louis County, 2015–present) | Unclear. County police performed police work in and around Ferguson, and there is overlap between unlawful |

¹⁴⁸ JEFFREY FAGAN ET AL., FINAL REPORT: AN ANALYSIS OF RACE AND ETHNICITY PATTERNS IN BOSTON POLICE DEPARTMENT FIELD INTERROGATION, OBSERVATION, FRISK, AND/OR SEARCH REPORTS (2015), <https://raceandpolicing.issuelab.org/resources/25203/25203.pdf>.

¹⁴⁹ INVESTIGATORY STOP AND PROTECTIVE PAT DOWN SETTLEMENT AGREEMENT, *supra* note 53.

¹⁵⁰ Chicago Memorandum of Agreement, *supra* note 60.

¹⁵¹ Consent Decree, *supra* note 61.

¹⁵² Consent Decree, *United States v. City of Ferguson*, No. 4:16-cv-000180-CDP (E.D. Mo. Mar. 17, 2016), <https://www.justice.gov/crt/file/833701/download>.

| | | | | | |
|----------------------|--|--|----------------------|---|---|
| | | | | | practices and the County's predictive policing use. |
| Miami (FL) | DOJ Memorandum of Agreement (2016) ¹⁵³ | Pattern and practice of excessive force with use of firearms | 2008–2011 | HunchLab (2014–2017), IBM (2011) | IBM system used historical crime data. |
| Maricopa County (AZ) | Federal District Court Order (2013), ¹⁵⁴ Federal District Court Order (2016), ¹⁵⁵ DOJ Settlement Agreement (2015) ¹⁵⁶ | Unconstitutional and racially biased stops, searches, and arrests; unconstitutional lengthening of stops; unlawful retaliation against people who made complaints or criticized MCSO | 2007–2011; 2014–2017 | PredPol (Mesa, 2016–present); HunchLab (Peoria, 2015–present); RTMDx (Glendale, 2012 pilot); BJA-funded pilot (Tempe, 2014) | Unclear. Mesa shares data with MCSO, and PredPol uses crime data. |
| Milwaukee (WI) | Settlement Agreement with Five-Year Consent Decree (2018) ¹⁵⁷ | Unconstitutional and racially biased stop and frisk practices; draft DOJ investigation report found racial disparities in traffic | Since 2008 | Milwaukee County received DOJ funds for data-driven policing. | Prior police chief expressed interest. Unclear if County predictive |

¹⁵³ AGREEMENT BETWEEN THE UNITED STATES DEPARTMENT OF JUSTICE AND THE CITY OF MIAMI REGARDING THE CITY OF MIAMI POLICE DEPARTMENT (2016), <https://www.justice.gov/crt/file/833286/download>.

¹⁵⁴ *Ortega Melendres v. Arpaio*, No. CV-07-02513-PHX-GMS, 2013 WL 5498218 (D. Ariz. Oct. 2, 2013), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-order>.

¹⁵⁵ *Ortega Melendres v. Arpaio*, No. CV-07-02513-PHX-GMS, 2016 WL 2783715 (D. Ariz. May 13, 2016), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-2016-order>.

¹⁵⁶ Settlement Agreement, *supra* note 104.

¹⁵⁷ Letter from the Office of the City Attorney to the Common Council of the City of Milwaukee Regarding Proposed Settlement in *Charles Collins, et al. v. City of Milwaukee, et al.*, Case No. 17-CV-0234-JPS (Apr. 27, 2018), https://city.milwaukee.gov/ImageLibrary/Groups/cityFPC/agendas5/180503EX_III.pdf.

| | | | | | |
|------------------|--|--|-----------|---|--|
| | | enforcement practices | | | policing system uses Milwaukee Police Department data. |
| New Orleans (LA) | DOJ Consent Decree (2013) ¹⁵⁸ | Pattern or practice of excessive use of force; unlawful stops, searches, and arrests; discrimination based on race, national origin, and LGBT status; gender discrimination in failure to investigate violence against women | 2005–2011 | Palantir (2012–2018) | Palantir used NOPD data to train and use the system. |
| New York (NY) | Remedial Process (2013), ¹⁵⁹ Settlement Agreement (2015) ¹⁶⁰ | Unconstitutional and racially biased stops, searches, and arrests | 2003–2012 | HunchLab (2015 pilot); NIJ-funded pilot (funded in 2013 for three-year grant); Palantir (unknown) | HunchLab used historical NYPD crime data. |
| Newark (NJ) | DOJ Consent Decree (2016) ¹⁶¹ | Pattern or practice of unlawful stops, searches, and arrests; racially biased policing practices; excessive | 2007–2012 | Starlight (unknown); PredPol (unknown); RTMDx (2012 pilot) | Pilot occurred during DOJ investigation, so it is possible |

¹⁵⁸ OFFICE OF THE CONSENT DECREE MONITOR NEW ORLEANS, LA., *supra* note 85.

¹⁵⁹ *Floyd v. City of New York*, 959 F. Supp. 2d 540 (S.D.N.Y. 2013).

¹⁶⁰ Stipulation of Settlement and Order, *Davis v. City of New York*, No. 10 Civ. 0699 (SAS) (S.D.N.Y. Feb. 4, 2015), <https://www.clearinghouse.net/chDocs/public/PN-NY-0013-0015.pdf>.

¹⁶¹ Consent Decree, *United States v. City of Newark*, No. 2:16-cv-01731-MCA-MAH (D.N.J. Apr. 29, 2016), <https://www.justice.gov/crt/file/868131/download>.

| | | | | | |
|-------------------|---|--|---------------------------|--|--|
| | | use of force; theft by Newark Police Department (NPD) officers | | | system used biased NPD data. PredPol uses NPD data on past crime, place/time of crime. |
| Philadelphia (PA) | Settlement Agreement and Consent Decree (2011) ¹⁶² | Unconstitutional and racially biased stop and frisk practices | 2005–2010; Jan.–June 2017 | BJA-funded pilot (2010); HunchLab (2013–present) | Philadelphia Police Department data used in pilot. Recent finding of biased practices can be reflected in data currently being used. |
| Seattle (WA) | DOJ Consent Decree (2012) ¹⁶³ | Pattern or practice of excessive use of force. DOJ did not issue a finding of racially biased policing but expressed serious concerns about bias particularly in pedestrian encounters | 2009–2011 | PredPol (2013–present) | Seattle Police Department documents state the PredPol system uses crime data dating back to 2008. |

¹⁶² Settlement Agreement, Class Certification, and Consent Decree, *Bailey v. City of Philadelphia*, No. 10-5952 (E.D. Pa. June 21, 2011), <https://www.clearinghouse.net/chDocs/public/PN-PA-0013-0002.pdf>.

¹⁶³ Settlement Agreement and Stipulated Order of Resolution, *United States v. City of Seattle*, No. 12-CV-1282 (W.D. Wash. July 27, 2012), <https://www.clearinghouse.net/chDocs/public/PN-WA-0001-0002.pdf>.

| | | | | | |
|------------------------|--|------------------------------------|-----------|--|--|
| Suffolk County (NY) | DOJ Memorandum of Agreement (2014) ¹⁶⁴ | Biased policing against Latinos | 2004–2011 | Vendor unknown, but DOJ and Foundation money were awarded to develop or acquire a system. | Media reports indicate use of a predictive policing system since 2012 but vendor unknown. |
|------------------------|--|------------------------------------|-----------|--|--|

¹⁶⁴ AGREEMENT BETWEEN THE UNITED STATES DEPARTMENT OF JUSTICE AND SUFFOLK COUNTY POLICE DEPARTMENT (2014), <https://www.clearinghouse.net/chDocs/public/PN-NY-0040-0003.pdf>.