Algorithmic Policing

Ranae Jabri*

Click here for the most recent version
November 2021

Abstract

Predictive policing algorithms are increasingly used by law enforcement agencies in the United States. These algorithms use past crime data to generate predictive policing boxes, specifically the highest crime risk areas where law enforcement is instructed to patrol every shift. I collect a novel dataset on predictive policing box locations, crime incidents, and arrests from a major urban jurisdiction where predictive policing is used. Using institutional features of the predictive policing policy, I isolate quasi-experimental variation to examine the causal impacts of algorithm-induced police presence. I find that algorithm-induced police presence decreases serious property and violent crime. At the same time, I also find disproportionate racial impacts on arrests for serious violent crimes as well as arrests in traffic incidents i.e. lower-level offenses where police have discretion. These results highlight that using predictive policing to target neighborhoods can generate a tradeoff between crime prevention and equity.

Keywords: algorithms, predictive policing, law enforcement, crime, neighborhoods, big data, inequality, race
JEL codes: K40, J15, K42, H0, C53

*Department of Economics, Duke University. Email: r.jabri@duke.edu. I am indebted to Patrick Bayer, Robert Garlick, Seth Sanders, and Christopher Timmins for their guidance, encouragement and support. Thank you to Bocar Ba, Jason Baron, Phil Cook, Dhammika Dharmapala, Jennifer Doleac, Erica Field, Matt Masten, Arnaud Maurel, John Rappaport, Juan Carlos Suarez-Serrato, Pengpeng Xiao and seminar participants at the Public Lab, Labor Lunch, and DevTea at Duke for helpful questions and comments. I also thank the jurisdictions that provided access to the data that made this paper possible, and PredPol for providing access to their site. All errors are my own.

1 Introduction

Predictive policing algorithms are increasingly used by law enforcement agencies in the hopes of improving efficiency. A leading predictive policing company, PredPol, claims to be "used to help protect one out of every 33 people in the United States."[1] Predictive policing algorithms use past crime data to predict crime risk for every geographic unit (e.g. 500 feet by 500 feet area squares) in a jurisdiction at high frequency, for every shift. The highest crime risk geographic units are designated as predictive policing boxes. Every shift, law enforcement receives electronic maps delineating these predictive policing boxes where they are instructed to patrol. While predictive policing is increasingly widespread, the impacts of neighborhood targeting brought about by predictive policing on crime and racial disparities in arrests are open questions.

The effect of local police presence is a longstanding empirical question. An established literature studies the effects of large-scale, long-term police deployments.[2] Yet, there is little evidence on the causal effects of local police presence,[3] and even less evidence on whether local police presence has disproportionate racial impacts. The locations where law enforcement patrol within jurisdictions are likely related to crime and arrest rates, as well as to racial composition of neighborhoods. This endogeneity complicates estimating the causal effects of local police presence. With the recent adoption of predictive policing, patrol is allocated in a systematic way based on the predicted crime risk score. There is high frequency data on where law enforcement is induced to patrol, and I can isolate areas with similar crime risk where law enforcement is not induced to patrol – for every shift in a jurisdiction. That is, I can compare areas with similar crime risk but different levels of police presence.

In this paper, I investigate the impacts of police presence induced by predictive policing algorithms on crime and racial disparities in arrests. In comparison to the previous literature,

[1] https://blog.predpol.com/predpol-named-to-govtech100-list-for-5th-straight-year

[2] Papers investigate increases in police hiring (Mello, 2019; Chalfin et al., 2020), large-scale city-wide deployments (Di Tella and Schargrodsky, 2004; Klick and Tabarrok, 2005; Draca et al., 2011), and long-term traditional hotspots (Weisburd and Telep, 2014; Blattman et al., 2021)

[3] Blanes i Vidal and Mastrobuoni (2018); Weisburd (2021)

my focus on local police presence addresses different policy questions: What happens to crime when law enforcement patrol a specific area? Are people of color disproportionately affected? I contribute to this literature by studying these questions at a much more granular level with respect to time and space than prior papers. To date, there has been no study that has considered both the overall efficacy and equity implications of the causal impacts of local police presence.

To examine these questions, I collect a unique dataset describing predictive policing box locations, crime incidents, and arrests from a major urban jurisdiction in the United States. The jurisdiction uses PredPol, one of the first predictive policing technologies to be deployed in the United States. I use institutional context details to isolate quasi-experimental variation in police presence induced by predictive policing algorithms. My research design exploits an exogenous change in the PredPol system in the set of PredPol boxes that are delivered to law enforcement. Before the change, PredPol generated its boxes using aggravated assault, auto burglary, motor vehicle theft, robbery, and shots fired violation codes. After the change, two more violation codes (residential and commercial burglary) were added to the original set.

Several features of the institutional setting and data make it possible to isolate quasi-experimental variation of algorithm-induced police presence from this exogenous change. First, after the change, law enforcement only sees the new sets of predictive policing boxes every shift, and the change is not salient to law enforcement. Second, the change happens to all districts of the jurisdiction at once, and is unlikely to be correlated with underlying district-level time-varying unobservables that could be driving crime and arrest outcomes. Third, I observe where predictive policing boxes for the original crimes would have been if the change had not happened, that would have been delivered to law enforcement if not for the change. This set of predictive policing boxes that are not delivered after the change have similar underlying crime risk and make up a control group for predictive policing boxes that are delivered to police before the change.

The empirical strategy compares outcomes at predictive policing boxes that are delivered to law enforcement before the change with outcomes at boxes that are not treated after the change as a result of this quasi-random change, accounting for the predictive policing boxes that are treated after the change and box fixed effects. Additionally, using institutional knowledge about how PredPol predicts crime risk along with data on input variables used to generate predictive policing boxes, I further account for any endogeneity concerns arising from the possibility that treatment and control groups have different underlying crime risk.[4]

When law enforcement patrol an area, the perceived probability of arrest is expected to increase, increasing the expected costs of committing a crime, and decreasing the likelihood that a crime is committed (Becker, 1957). At the same time, law enforcement has discretion to stop civilians,[5] and traffic stops are the most common reasons for contact with the police.[6] Lower-level offenses that may have been previously undetected may be more likely to be reported as crimes. Police also have discretion to make arrests for lower-level offenses.[7] Overall, I find that algorithm-induced police presence decreases serious property and violent crime on the order of 3 crimes per 1000 PredPol boxes, and do not find evidence of crime displacement to surrounding boxes. I also find suggestive evidence that reported traffic incidents increase.

Next, I test if algorithm-induced police presence has racially disparate impacts on arrests. I use a nested model building on the main empirical strategy to estimate effects of algorithm-induced police presence for Black, Hispanic and white individuals. Using the nested model, I calculate the counterfactual mean of arrests for predictive policing boxes if they had not been delivered as predictive policing boxes and treated. Then, to test if

[4] I proxy for PredPol's underlying crime risk measure using box fixed effects and crime lags included in prediction for both sets of predictive policing boxes. Robustness specifications also control for district-time fixed effects and time trends to account for any time-varying unobservables that could be driving outcomes.

[5] Justice Sotomayor in Utah v. Strieff, 2015: "This Court has allowed an officer to stop you for whatever reason he wants—so long as he can point to a pretextual justification after the fact. Whren v. United States, 517 U. S. 806, 813 (1996). "

[6] Bureau of Justice Statistics. https://www.bjs.gov/index.cfm?tid=702&ty=tp

[7] Justice Sotomayor in Utah v. Strieff, 2015: "The officer's control over you does not end with the stop. If the officer chooses, he may handcuff you and take you to jail for doing nothing more than speeding, jaywalking, or "driving [your] pickup truck . . . with [your] 3-year-old son and 5-year-old daughter . . . without [your] seatbelt fastened." Atwater v. Lago Vista, 532 U. S. 318, 323–324 (2001)."

algorithm-induced police presence has racially disparate impacts on arrests, I compare the estimates *by race* of the effects weighted by the counterfactual mean of arrests that would have happened if PredPol boxes had not been treated as predictive policing boxes. I find evidence that algorithm-induced police presence has disproportionate racial impacts on arrests for serious violent crimes and arrests in traffic incidents for Black individuals compared to white individuals. For arrests for serious violent crimes, this disparate impact is driven by a statistically significant decrease in arrests of white individuals. For traffic incidents, I find suggestive evidence that the disparate impact is driven by more arrests in traffic incident arrests of Black individuals. Overall, these results imply that while there is evidence that algorithm-induced police presence deters crime, there is also evidence that racial disparities in arrest rates are increased, raising serious equity concerns about neighborhood targeting that results from predictive policing.

As a robustness check, I use a Regression Discontinuity Design framework to estimate the effect of algorithm-induced police presence. PredPol predicts a continuous measure of underlying crime risk, and there are algorithmic cutoffs for a box to be a PredPol box. This alternative empirical strategy uses the discontinuities in PredPol box treatment at the underlying crime risk algorithmic threshold. While I do not observe the continuous crime risk measure, I use institutional knowledge from PredPol's marketing materials and a publication by authors affiliated with PredPol (Mohler et al., 2015) along with input data used to generate predictive policing boxes, to predict an estimate of the continuous crime risk measure using a machine learning model. I use this estimate of the continuous crime risk measure as a running score in a regression discontinuity design framework (Boehnke and Bonaldi, 2019). The framework compares outcomes at predictive policing boxes that marginally make the threshold to be a PredPol box with boxes that marginally miss the threshold to be designated as a PredPol box. Using this Regression Discontinuity Design framework, I also find that algorithm-induced police presence decreases serious property and violent crime.

This paper makes contributions to several strands of research. Experiments study the effects of using predictive policing to allocate patrols compared to other approaches at the district-level, examining effects on crime (Mohler et al., 2015; Ratcliffe et al., 2020) and racial disparities in arrests (Brantingham et al., 2018). Compared to these papers, my paper is a much more granular study of what actually happens at predictive policing boxes, including who is affected.

This paper also contributes to several different lines of the literature on the effects of police presence on crime, as mentioned earlier.[8] There is less evidence on the effect of local short-term police presence on crime (Weisburd, 2021; Blanes i Vidal and Mastrobuoni, 2018). I contribute to the literature by estimating the effects of local police presence at a more granular level with respect to time and space. My paper is also closely related to papers on the effects of traditional hotspots (Weisburd and Telep, 2014; Braga et al., 2019; Blattman et al., 2021).[9] Compared to predictive policing boxes that change every shift and have uncertain locations, traditional hotspots stay fixed over time, with certain patrol locations, and crime may simply be displaced.[10] To my knowledge, the literature on traditional hotspots does not study the effects of hotspots on racial disparities in arrests.

It is an open question whether local police presence has disproportionate racial impacts. A recent paper Chalfin et al. (2020) finds that effects of large police forces differ by race. Police force size causally increases the number of low-level "quality-of-life" offenses like drug possession and disorderly conduct, in particular for Black individuals. Analogous to Chalfin et al. (2020), I suggestive evidence that algorithm induced police presence increases the number of traffic incidents, and find the impact of algorithm-induced police presence is disproportionately larger for Black arrestees compared to white arrestees in traffic incidents – where police have more discretion. Chalfin et al. (2020) also finds that larger police forces

[8]See Durlauf and Nagin (2011); Chalfin and McCrary (2017) for a review and papers referenced earlier.

[9]See Braga et al. (2019) for a review.

[10]When traditional hotspots remain fixed over time, potential offenders may be aware of the location of police presence and may displace to high crime risk control areas. If it is costly to displace when criminals observe police presence, estimated effects of predictive policing boxes may be less affected by displacement effects than traditional hotspots.

decrease serious property and violent crimes, in particular for Black suspects. While I also find that serious property and violent crime decrease, I find that the impact of algorithm-induced police presence is disproportionately larger for Black arrestees than white arrestees. My findings suggest that increased local police presence has disproportionate racial impacts on Black individuals compared to white individuals.

Predictive policing is an example of algorithms used in decision making. Algorithms are becoming ubiquitous in society, and this paper contributes to the emerging literature studying algorithms used in decision making (Lum and Isaac, 2016; Kleinberg et al., 2017, 2018; Stevenson and Doleac, 2019; Mastrobuoni, 2020). Predictive models are considered to be contentious because of concerns that using past data to predict future risk may amplify pre-existing racial inequities (O'Neil, 2017). PredPol claims to not explicitly use race in models; however, past data may reflect historical patterns and biases in policing, and models may replicate these disparities (Lum and Isaac, 2016). Moreover, these disparities may be amplified if police discover more crime at predictive policing boxes which can create a negative feedback loop (Lum and Isaac, 2016).

My finding – that algorithm-induced police presence has impacts on crime – provides evidence of this feedback, that could amplify any pre-existing inequities through algorithmic feedback loops. The sign of the estimated effects of algorithm-induced police presence differ by crime types, revealing that the direction of the feedback loop depends on the crime types used to predict predictive policing boxes. Taken together with concerns that predictive policing box locations can reflect historical disparities policing (Lum and Isaac, 2016), my findings imply that an important equity consideration is to carefully consider the types of crimes being included in prediction.

The roadmap of the paper is as follows: Section 2 provides an overview of the context that I study where the predictive policing tool, PredPol, is used. Section 3 describes the novel data used for my analysis. Section 4 outlines the quasi-experimental research design that I use to identify the effects of predictive policing on crime incidents, and Section 4.2

presents results. Section 5 extends the main quasi-experimental research design to test for disproportionate racial impacts of algorithm-induced police presence on arrests, and Section 5.3 discusses results. Section 6 presents the Synthetic Regression Discontinuity framework of Boehnke and Bonaldi (2019) applied to the predictive policing context, as a robustness for my main quasi-experimental research design, and Section 6.3 presents results. Finally, Section 7 concludes and discusses policy implications.

2 Context

In this section, I present the jurisdiction, predictive policing technology (PredPol), and exogenous change that I use in my main empirical strategy to examine the effects of algorithm-induced police presence. All information about PredPol and its software comes from public information that I describe. Information on how PredPol is used by the jurisdiction that I study comes from public sources or conversations with decision-makers in the jurisdiction.

2.1 Jurisdiction

I study a major urban jurisdiction with a population of over 1 million people, contained in a metropolitan statistical area among the fifty largest in the United States.[11] The race/ethnicity population breakdown is around 15-20% Non-Hispanic Black, 40-45% Non-Hispanic white, and 25-30% Hispanic. A large law enforcement agency using predictive policing serves the jurisdiction, with over 2000 sworn officers and civilian employees. The jurisdiction has a uniform patrol division that is assigned at the district and shift level, who patrol in law enforcement uniforms.

[11] I promised the jurisdiction that I would not reveal the name of their jurisdiction.

2.2 PredPol

PredPol was founded in 2012, and is one of the first predictive policing companies to be used in the United States.[12] As of 2015, almost sixty US cities used PredPol.[13] PredPol boxes are delivered to law enforcement through an online interface and also patrol reports. Figure 1 shows an example of a patrol report from a PredPol guide. PredPol instructs officers to go to PredPol boxes for "about 6 minutes per hour".[14] In practice, in the jurisdiction that I study, patrols are instructed to go to PredPol boxes in their down time between calls for service. Patrols are also instructed to patrol in the PredPol boxes as they would normally patrol.

PredPol splits jurisdictions into 500 feet by 500 feet boxes, and then predicts a continuous crime risk measure for all boxes every shift (Mohler et al., 2015). According to PredPol marketing and a publication of coauthors affiliated with PredPol (Mohler et al., 2015), PredPol only uses crime time, crime type and GPS to predict crime risk for a set of crime types. Districts are split into 500ft by 500ft boxes, then PredPol predicts continuous crime risk for all boxes in every shift. Crime risk is predicted as an exponential decay function of the crime lags in each box and a crime time-invariant box effect. The functional form of the crime risk probabilistic rate (λ_{it}) of events in box i at time t (Mohler et al., 2015) is as follows:

$$\lambda_{it} = \mu_i + \sum_{t_i^n < t} \theta \omega e^{-\omega(t - t_i^n)} \tag{1}$$

where t_i^n are times of events in box i in the history of the process in the window being used for prediction T, T is suggested to be 365 days, μ_i is a baseline Poisson process rate (constant long-term background rate) or time-invariant box parameter, and $\theta \omega e^{\omega(t - t_i^n)}$ is an exponential decay "contagion" effect in crime data to capture short-term dynamics.

A fixed number of boxes per district and shift are designated as PredPol boxes, and the

[12] https://blog.predpol.com/predpol-named-to-govtech100-list-for-5th-straight-year
[13] https://www.forbes.com/sites/ellenhuet/2015/02/11/predpol-predictive-policing/#4731ae3e4f9b
[14] https://www.predpol.com/law-enforcement/#predPolicing

24 boxes with the highest crime risk per district per shift are designated as PredPol boxes. A box i in district d at time t is a PredPol box if the crime risk for the box at time t (λ_{it}) is greater than or equal to the 24th highest risk box for the district ($max^{24}_{dt}\{\lambda_{1t}, \ldots, \lambda_{It}\}$).[15]

2.3 Exogenous change in crimes used to predict day shift predictive policing boxes

PredPol boxes for "All Crimes" are delivered to law enforcement patrol units during the day shift in all districts of the jurisdiction. Prior to 11/20/2019, auto burglary offenses, vehicle theft offenses, robbery offenses, assault offenses and shots fired calls for service are the crime types used to generate "All Crimes" PredPol boxes. After 11/20/2019, two more crime types – residential burglary and commercial burglary– are added to the set of crime types used to generate PredPol boxes for "All Crimes", and the PredPol boxes for the new expanded crime types are delivered to law enforcement. Henceforth, I will refer to the PredPol boxes predicted using the original set of crimes types as All Crimes PredPol boxes, and PredPol boxes predicted using the new set of crime types adding the two additional crime types as All Crimes Plus PredPol boxes. Around a window of the exogenous change on 11/20/2019[16], the daily number of All Crimes and All Crimes Plus crime types during the day shift are highly correlated, with a correlation of 0.92.

I interviewed the key law enforcement decision maker in charge of PredPol in the jurisdiction. According to the law enforcement decision maker, the change came about randomly. The change happens to all districts of the jurisdiction at once, and is unlikely to be correlated with underlying time-varying unobservables at the district-level that could be driving crime or arrest outcomes. Moreover, this change was not salient to law enforcement who only saw that the PredPol boxes were for "All Crimes" PredPol boxes in the patrol reports. Without the change, law enforcement would have continued to receive the All Crimes PredPol boxes

[15] $PredPolBox_{idt} = 1(\lambda_{idt} \geq max^{24}_{dt}\{\lambda_{1t}, \ldots, \lambda_{It}\})$

[16] 5/20/2019-3/1/2020, which I use for the quasi-experimental empirical strategy estimation window, exploiting this exogenous change

for the original crime types, and after the change, law enforcement received the All Crimes Plus PredPol boxes for the original crime types plus two more crime types. A nice feature of the institutional setting that helps to identify treatment effects is that PredPol kept predicting All Crimes PredPol boxes in the background even after they were no longer delivered to law enforcement.

3 Data

Estimating the effect of algorithm-induced police presence on crime incidents and racial disparities in arrests requires detailed data on predictive policing box locations, crime outcomes, arrests, and race of arrestees. To conduct my analysis, I assemble a unique data set that makes this analysis possible, combining novel data on (1) predictive policing box locations, (2) crime incident/calls for service data for crime types used in PredPol crime prediction (input data to predict PredPol boxes), and (3) incident and arrest data from the jurisdiction. I collect PredPol box data which includes the location of PredPol boxes every shift in every district, and the crime types for which crime risk and PredPol boxes are predicted. This includes PredPol boxes for crime types that are actually delivered to law enforcement and PredPol boxes for crime types that are no longer delivered to law enforcement. My paper leverages the quasi-experimental variation in PredPol boxes that are actually delivered to police to circumvent issues of endogeneity that predictive policing boxes are located in high crime risk areas.

Data on actual police presence requires access to automobile locator systems (AVL) that more and more police departments are starting to use and track. Unfortunately, the jurisdiction that I study does not record this kind of data.[17] While I do not have data on police presence itself, I use the predictive policing policy instrumentally to identify the effect of police presence using an intent to treat analysis.

[17]Use of such data is rare, though Weisburd (2021) and Blanes i Vidal and Mastrobuoni (2018) use AVL data to study the effect of police presence on overall crime outcomes, identifying effects using plausibly exogenous shifts in police presence.

I describe each novel source of data that I collect. Then, I detail how I assemble the panel dataset at the box-shift level that I use to conduct my analysis, and in particular how I map crime incidents and arrests to the box and shift in which they occured:

1. **PredPol box locations data:** I observe locations of 500 feet by 500 feet PredPol boxes (GPS coordinates) and the shift/date for which they are generated. I observe the location of All Crimes PredPol boxes that PredPol predicts in the background even after they are no longer delivered.[18] I create a list of the locations of the All Crimes and All Crimes Plus PredPol boxes over a three year period (3/1/2018-3/1/2020). I call this list the "ever-PredPol boxes."

2. **Crime offense/calls for service data for crime types used in PredPol box prediction:** For each offense, I observe the date, start/end time, offense types, address and GPS location. I also see whether the offense is excluded from prediction. Offenses can be excluded because they have a long duration (start to end time), the offense does not properly geocode, or if they are a duplicate. I map offenses to the boxes in which they occur using the GPS.

3. **Incident/arrest data from jurisdiction:** The jurisdiction provides incident and arrest-level data for my analysis. For each incident that results in an incident report, I observe the incident nature, incident report date and time, address (which I geocode using the *Google Maps Application Programming Interface*), suspect race, and victim race. For each incident, I observe the arrests that occured. If the arrest is able to be physically made at the time of the incident report, I observe the arrest date/time and address. There are also arrests that happen after the incident report time for which I observe the arrest date/time and address. Because of these unresolved data complications, I am not currently using the arrest date/time or address for my analysis. For arrests, I observe the race/ethnicity, age, gender of the arrestee.

[18] Unfortunately, I only observe All Crimes Plus PredPol boxes a few days before they are delivered.

Panel dataset creation: Using the list of "ever-PredPol boxes", I create a panel data set of box-time (date/shift) observations over a three year period (3/1/2018-3/1/2020). A box i is included if it is ever a PredPol box over the three year period. For every box-time observation, I observe the PredPol box treatment status of box i at time t.[19] The outcomes of interest are crime incidence in box i at time t, [20] arrests of Black individuals in box i at time t, arrests of Hispanic individuals in box i at time t, and arrests of white individuals in box i at time t[21]

3.1 Descriptive Statistics

There are 8,224 ever-PredPol boxes in the three-year ever-PredPol box panel dataset that I assemble. These boxes will be either All Crimes or All Crimes Plus PredPol boxes at least once over the three year period from 3/1/2018 to 3/1/2020. 1,924 boxes will be an All Crimes PredPol boxes at least once over the window (5/20/2019-3/1/2020) around the exogenous change in the crimes used to predict PredPol boxes, which I use for the quasi-experimental empirical strategy estimation window. Figure 3 plots the distribution of the number of day shifts a box will be an All Crimes PredPol boxes from the sample of the boxes that will ever be an All Crimes PredPol boxes at least once over the window 5/20/2019-3/1/2020. Of

[19] I observe the All Crimes PredPol boxes that are not treated after 11/20/2019.

[20] I map crime incidents using the incident address to boxes in which they occur. For crime types used for prediction (auto burglary offenses, vehicle theft offenses, robbery offenses, residential burglary offenses, commercial burglary offenses, assault offenses and shots fired calls for service), I observe the incident start time, which I use to map incidents to the shift in which they started. For incidents for crime types outside of this set of crime types, I only observe the incident report time, which I use to map these incidents to the shift in which they occured. For the subset of crime offenses used in prediction (contained in both the second and third sources of data above), I examine the difference between the incident start time and the incident report time. The difference is large for burglary and motor vehicle theft offenses. Burglary and motor vehicle theft are non-violent property crimes that do not involve a personal threat of violence. Incidents might be reported at a later time after discovery. There is a smaller difference between incident and report time for assault and robbery offenses, which are violent crimes where there is a personal threat of violence. Based on this analysis, I exclude crime types from my analysis where there may be gaps between incident time and incident report time. In the future, I hope to get access to better incident start time data to perform analysis for more types of crime incidents. For example, I include traffic incidents where law enforcement are likely initiating traffic stops, and the incident start time. There is unlikely to be much time lapsed to the incident report time.

[21] I map arrests using the incident address and incident time. For arrests for incidents among the crime types used for prediction, I use the incident start time, and for arrests for all other incidents, I use the incident report time for the jurisdiction data.

these boxes, 47.56% will be PredPol boxes for less than six day shifts. 83.06% of the boxes will be PredPol boxes for thirty or fewer day shifts. 93.61% boxes will be PredPol boxes for 100 or fewer day shifts. This paper uses this high-frequency variation of boxes switching in and out of being PredPol boxes every shift.

4 Main Empirical Strategy: the effects of algorithm-induced police presence

What are the effects of predictive policing boxes on crime outcomes? There are empirical complications to estimating the effects of algorithm-induced police presence: predictive policing box locations are located in the highest predicted crime risk areas within a district. Directly estimating the effect of algorithm-induced police presence, without accounting for the underlying crime risk of boxes, would lead to omitted variable bias. To circumvent these endogeneity concerns, I use data on boxes that are nearly predictive policing boxes with similar crime risk and context to isolate quasi-experimental variation in police presence induced by predictive policing algorithms.

I leverage the addition of two crime types to the original crime types used to predict All Crimes PredPol boxes that changes the set of PredPol boxes delivered to police, and the data that I observe on predictive policing box locations that are not delivered or active after the change to estimate treatment effects. Figure 4 illustrates the research design. Prior to 11/20/2019, law enforcement receives the All Crimes PredPol boxes in blue, and these PredPol boxes are "active" and treated. After 11/20/2019, PredPol kept predicting All Crimes PredPol boxes even when they were no longer delivered or "active", and they serve as a control group in my research design. After 11/20/2019, law enforcement receives the All Crimes Plus PredPol boxes in yellow which are "active" and treated. There is also overlap between the All Crimes and All Crimes Plus PredPol boxes since there is overlap in the crime types used in prediction. My research design compares the active All Crimes PredPol boxes to the inactive All Crimes Plus PredPol boxes, accounting for the overlap between the All Crimes and All Crimes PredPol boxes and the fact that after the change, the All Crimes Plus PredPol boxes are active.

To estimate the effect of treated predictive policing boxes, and the effect of algorithm-induced police presence, I estimate the following model using a window around the exogenous change from 5/20/2019 to 3/1/2020[22]:

$$\begin{aligned} Y_{it} =& \beta Active\_AC_{it} + \delta AC_{it} + \xi Active\_ACPlus_{it} \\ &+ \sum_{j=1}^{T} \gamma_j y_{it-j} + \mu_i + \phi_{dt} + \varepsilon_{it} \end{aligned} \tag{2}$$

where Y_{it} are crime incidents for box i at time t, $Active\_AC_{it}$ is an indicator for whether box i is an active All Crimes PredPol box at time t, AC_{it} is an indicator for whether box i is an All Crimes PredPol boxes at time t (active or inactive), $Active\_ACPlus_{it}$ is an indicator for whether box i is an active All Crimes Plus PredPol box at time t. $Active\_ACPlus_{it}$ is included to account for which boxes are actually active after the switch and any potential effect on crime, and the overlap between All Crimes and All Crimes Plus PredPol boxes. y_{it-j} are crime lags summing crimes included in PredPol box prediction for both All Crimes PredPol boxes and crime lags summing crimes included in PredPol box prediction for All Crimes Plus PredPol boxes. I also include box fixed effects μ_i to account for any time-invariant box characteristics. β is the effect of algorithm-induced police presence at predictive policing boxes.

I address any additional endogeneity concerns that All Crimes and All Crimes Plus PredPol boxes have different time-varying underlying crime risk by including a rich set of controls. I include controls for PredPol's underlying crime risk λ_{it} using box fixed effects and crime lags to address concerns that time-varying underlying crime risk (omitted variable) could be correlated with ε_{it}. As a robustness check, I also include district-time fixed effects/trends

[22] I cutoff of the sample in March 1, 2020 around the start of the Covid-19 pandemic.

ϕ_{dt} to control for any unobservable time-varying district level trends that could be driving the change or outcomes. Standard errors are clustered at the box level.

The identifying assumption is that conditional on a box i being an All Crimes PredPol box at time t, a box i being an Active All Crimes Plus PredPol box at time t, proxying for PredPol's underlying crime risk (box fixed effects and crime lags for both All Crimes and All Crimes Plus PredPol boxes), and district-time fixed effects/trends, a box i at time t being an Active All Crimes PredPol boxes must be orthogonal to omitted variables that could also affect crime in box i:

$$E[Active\_AC_{it} \cdot \varepsilon_{it} | AC_{it}, Active\_ACPlus_{it}, \mu_i, y_{it-j}, \phi_{dt} \ldots] = 0 \tag{3}$$

4.1 How can algorithm-induced police presence affect reported crime?

Next, I briefly discuss mechanisms for how predictive policing boxes and presence can affect reported crime incidents. First, police presence at predictive policing boxes may deter crime (Becker, 1957). Police presence increases the probability of apprehension and increases criminals' expected cost of committing a crime, which should decrease the likelihood that a crime is committed.

Second, there may be more police-civilian interactions at PredPol boxes as a result of patrol being targeted there. "Crime" incidents that would have been previously undetected may be discovered with increased police presence, which would increase reported crimes. This mechanism is more likely to affect crimes where police have to be there to discover a crime or have discretion. On the other hand, this mechanism is also less likely to affect property crimes, where people report stolen items themselves.

4.2 Results

This section presents estimates of algorithm-induced police presence on overall crime incidents and calls for service, which is β from equation 2. First, I examine the effects of algorithm-induced police presence on shots fired calls for service in Table 1. In the data, these calls for service are unlikely to become crime incidents. Column (1) shows the baseline specification estimate which additionally controls for box fixed effects and crime lags to address any additional endogeneity concerns about the quasi-experiment. In particular, I use box fixed effects and crime lags to proxy for PredPol's prediction of underlying crime risk. As a further check on robustness, Column (2) shows the estimate for the baseline model additionally controlling for district-time trends, and Column (3) shows the estimate for the baseline model additionally controlling for district-time fixed effects. I find a statistically significant increase in the reports of shots fired calls for service in the range of 0.513-0.578 calls per 1000 boxes – a large increase compared to the PredPol box outcome mean of 0.486, and that result is robust across specifications. While I do not directly observe patrol location and presence, these estimates validate that patrols are spending time in predictive policing boxes locations. Shots fired reports could be made by either civilians or law enforcement. Civilians may be more likely to report shots fired if law enforcement is nearby and patrol could be in the PredPol boxes and hear shots fired that they call in. More shots fired calls for service also imply that law enforcement are responding to calls for service in PredPol box areas, increasing activity.

Second, I examine the effect of algorithm-induced police presence on serious property and violent crime incidents which are aggravated assault, burglary (auto, commercial, residential), robbery and vehicle theft in Table 2. I find that algorithm-induced police presence statistically significantly deters around 2.909 to 3.077 crimes per thousand boxes, a nearly 30% reduction over the PredPol box mean of 9.348 crimes per thousand boxes. The finding is robust across specifications. As in Table 1, Column (1) shows the baseline specification estimate which controls for box fixed effects and crime lags to proxy for PredPol's prediction of underlying crime risk; Column (2) shows the estimate for the baseline model additionally controlling for district-time trends; and Column (3) shows the estimate for the baseline model additionally controlling for district-time fixed effects. As in Table 1, estimates are robust across specifications, and subsequent analysis of other outcomes will utilize the augmented specification in Column (3). The estimates are consistent with prior papers that find that police presence deters crime (Weisburd, 2021). Column (4) shows the effects of algorithm-induced police presence on serious property and violent crime incidents in predictive policing boxes and the 8 boxes that are adjacent, defined as the expanded box; the estimate is from the baseline specification augmented with district-time fixed effects. I expand the outcome to additionally include surrounding boxes to study whether crime displaces to nearby boxes. The effect algorithm-induced police presence into the expanded box area is a decrease of 4.844 crimes per 1000 boxes, a 19% reduction over the PredPol expanded box outcome mean of 25.325 crimes per 1000 expanded boxes –evidence that crime is not being displaced to the areas around PredPol boxes. Increased algorithm-induced police presence deters serious property and violent crime incidents in the expanded box.

Finally, I estimate the effect of algorithm-induced police presence on traffic incidents in Table 3. Column (1) uses the baseline specification additionally controlling for district-time fixed effects. I find suggestive evidence that algorithm-induced police presence increases the number of traffic incidents. While this effect is only marginally statistically significant, this finding is intuitive as traffic incidents involve lower-level offenses that are unlikely to be detected without law enforcement. When a patrol enters a predictive policing box, police-civilian interactions may increase, including traffic stops, which can result in traffic incidents.

Overall, these findings show that predictive policing deters serious property and violent crime, but there is also suggestive evidence that predictive policing increases traffic incidents, an example of a lower-level offense that is unlikely to be detected without police presence. Taken together, these findings suggest that the effects of algorithm-induced police presence differ by crime type.

5 Testing for disproportionate racial impacts

In this section, I examine whether algorithm-induced police presence at predictive policing boxes has disproportionate racial impacts. To test this, I use a nested model building on the main empirical strategy:

$$
\begin{aligned}
y_{it}^{R} = \sum_{r=b,h,w} 1\{R=r\}(&\beta_r Active\_AC_{it} + \delta_r AC_{it} + \xi_r Active\_ACPlus_{it} \\
&+ \sum_{j=1}^{T} \gamma_{j,r} y_{it-j} + \mu_{i,r} + \phi_{dt,r}) + \varepsilon_{it} \qquad (4)
\end{aligned}
$$

where y_{it}^{r} measures the number of arrests in box i at time t of individuals of race R; r takes values b (Black), h (Hispanic) , w (white). $Active\_AC_{it}$ is an indicator for box i being an All Crimes active PredPol box at t, AC_{it} is an indicator that box i is All Crimes PredPol box at t, and $Active\_ACPlus_{it}$ is an indicator that box i is an All Crimes Plus active PredPol box at t. β_r is the effect of police presence on arrests for individuals of race r.

5.1 Inframarginal and marginal arrests

The observed mean of arrests of race r at All Crimes active PredPol boxes:

$$
y_{obs,r} = \beta_r + \delta_r + \sum_{j=1}^{T} \gamma_{j,r} y_{it-j} + \mu_{i,r} + \phi_{dt,r} \qquad (5)
$$

The counterfactual mean of arrests of race r at All Crimes active PredPol boxes if treatment had not occured:

$$
\begin{aligned}
y_{cf,r} &= \delta_r + \sum_{j=1}^{T} \gamma_{j,r} y_{it-j} + \mu_{i,r} + \phi_{dt,r} \\
&= y_{obs,r} - \beta_r \qquad (6)
\end{aligned}
$$

The arrests of race r that are marginal because of police presence at All Crimes active PredPol boxes is the difference between $y_{obs,r}$ and $y_{cf,r} = \beta_r$. Therefore, the counterfactual arrest mean in the absence of treatment is the number of inframarginal arrests.

5.2 Testing if police presence have disproportionate racial impacts on arrests

To test if marginal arrests are disproportionately made up of Black arrests, I perform the two-sided test comparing the Black with the white marginal effects, weighted by how many arrests would have happened without the PredPol box (inframarginal arrests):

$$\frac{\textit{Black marginal arrests}}{\textit{Black inframarginal arrests}} > \frac{\textit{White marginal arrests}}{\textit{White inframarginal arrests}}$$
$$\frac{\beta_{black}}{y_{cf,black}} > \frac{\beta_{white}}{y_{cf,white}} \tag{7}$$

5.3 Results

I present results for tests for disproportionate racial impacts of algorithm-induced police presence on arrests of serious violent crime (aggravated assault, robbery) in Figure 5 and Table 4. Using estimates for marginal and inframarginal arrests from equation 4 and equation 6, I perform the two-sided test using equation 7. The left panel shows estimates for predictive policing boxes, and the right panel shows estimates for predictive policing boxes and surrounding boxes (expanded box outcome). I find that arrests for serious violent crime due to algorithm-induced police presence in predictive policing boxes and surrounding boxes are disproportionately likely to be Black arrestees. The left set of bars per panel plots the estimates of the number of marginal arrests by race; the center set of bars per panel plots the number of inframarginal arrests by race; the right set of bars per panel plots the marginal arrests weighted by the inframarginal arrests by race.

First, I find that algorithm-induced police presence has a statistically significant negative effect on the number of white arrests for both boxes and expanded boxes. The estimates for the Black marginal arrests and hispanic marginal arrest estimates are statistically insignificant for both the box and expanded box outcome. Second, there are nearly double the number of Black arrests than white arrests in the absence of treatment for both boxes and expanded boxes. Finally, I examine the marginal arrests weighted by the inframarginal arrests by race, to test whether arrests due to algorithm-induced police presence are disproportionately made up of arrests of people of color, compared to how many would have happened if there had not been a predictive policing box there.

For the box outcome, I find suggestive evidence that is marginally statistically significant at the 10% level that there are racial disparities in arrests for serious violent crime. The marginal arrests weighted by the inframarginal arrests is negative for white arrestees. Yet, there is no proportional decrease for Black arrestees, and the marginal arrests weighted by the inframarginal arrests is negative and close to zero for Black arrestees. For the expanded box outcome in the right panel, there is a similar fall in the marginal arrests weighted by the inframarginal arrests for white arrestees, and an increase in the number of arrests for Black arrestees. Moreover, I find statistically significant racial disparities in the impacts of algorithm-induced police presence, with a p-value of 0.019. For arrests of Hispanic arrestees: there is a small number of Hispanic inframarginal arrests, and I cannot conclude whether there are racial disparities in the impacts of algorithm-induced police presence compared to white individuals. Arrests due to algorithm-induced police presence in predictive policing boxes and surrounding boxes are disproportionately Black, evidence that there are disproportionate racial impacts of algorithm-induced police presence on arrests for serious violent crime.

Next, I fail to reject that there are racial disparities in the impacts of algorithm-induced police presence on arrests of serious property crime (burglary, vehicle theft) in Table 5. Examining estimates of marginal arrests by race, I find suggestive evidence that is statistically significant at the 10% level that algorithm-induced police presence decreases the number of

Black and white arrests for serious property crime. It is worth noting that the number of inframarginal Black arrests is over double the number of white arrests, and over three times the number of Hispanic arrests.

I examine whether there are racial disparities in the impacts of algorithm-induced police presence on arrests in traffic incidents in Figure 6. Law enforcement have greater discretion in traffic incidents than in other types of crimes. I find that arrests for traffic incidents due to algorithm-induced police presence in predictive policing boxes are disproportionately likely to be Black arrestees. The left set of bars per panel reports the estimates of the number of marginal arrests by race; the center set of bars per panel reports the number of inframarginal arrests by race; the right set of bars per panel reports the marginal arrests weighted by the inframarginal arrests by race. First, algorithm-induced police presence at predictive policing boxes have a marginally statistically significant positive effect on the number of Black arrests, and I find insignificant effects for Hispanic and white arrests. Ultimately, after accounting for how many traffic incidents would have happened without predictive police presence, there is evidence of racial disparities in the effects of algorithm-induced police presence, with Black arrestees disproportionately likely to be arrested for traffic incidents in predictive policing boxes.

6 Synthetic regression discontinuity design

In this section, I apply the synthetic regression discontinuity (SRDD) design framework of Boehnke and Bonaldi (2019) to further examine causal impacts of algorithm-induced police presence, as a robustness check to my main empirical strategy presented in Section 4. This framework uses the discontinuity in PredPol box treatment status at the algorithmic threshold to estimate treatment effects. Using institutional knowledge about how PredPol predicts underlying crime risk (λ_{it}) along with the algorithm input data, and the PredPol box treatment status, I predict an estimate of λ_{it}. The estimate of λ_{it} is a synthetic running score $\hat{\lambda}_{it}$, which I use as a running variable in a sharp regression discontinuity design.

Boehnke and Bonaldi (2019) provides a two-stage framework to identify the local average treatment effect using regression discontinuity design when the running variable is unobservable but treatment status is known. In the first stage, the synthetic score is predicted using treatment status, and the second stage uses the synthetic score as a running variable in RDD conditional on treatment status. The framework actually does not require that there is a continuous running score that is explicitly calculated by the decisionmaker that underlying treatment assignment, just that the treatment assignment "can be described as if it were implicitly based on such a score" (Boehnke and Bonaldi, 2019). Therefore, the PredPol context is a relevant setting to apply this framework, as I observe PredPol box treatment status which is explicitly based on the underlying crime risk prediction λ_{it}, which I do not observe (Mohler et al., 2015).

The framework drops misclassified boxes to guarantee a discontinuity in probability of PredPol box treatment AC_{it} at τ, the threshold of $\hat{\lambda}_{it}$:

$$\beta = \lim_{q \downarrow \tau} E[Y|\hat{\lambda} = q, AC_{it} = 1] - \lim_{q \uparrow \tau} E[Y|\hat{\lambda} = q, AC_{it} = 0] \tag{8}$$

The identifying assumptions are:

1. Continuity and smoothness of unobserved running variable λ_{it} and synthetic score $\hat{\lambda}_{it}$

2. Synthetic score perfectly predicts treatment status in first stage

In theory, the identifying assumption that the synthetic score perfectly predicts treatment status in the first stage is fulfilled in the PredPol context based on institutional knowledge. PredPol maintains that only three data points – crime type, crime time/date, crime GPS – are used to predict crime risk and PredPol boxes. Therefore, in theory, these three data points should perfectly predict crime risk and PredPol box treatment status. Next, I apply the two-stage framework of Boehnke and Bonaldi (2019) to examine the effects of algorithm-induced police presence in the PredPol setting:

6.1 First stage: estimating the underlying crime risk synthetic running score

In the first stage, I estimate the continuous score $\hat{\lambda}_{it}$ underlying PredPol box treatment status:

$$AC_{it} = h(i, d, y_{it-1}, \ldots) \tag{9}$$

where AC_{it} is an indicator variable for whether box i is an All Crimes PredPol box. I use a multilayer perceptron (MLP) neural network to predict whether a box i at time t is a PredPol box using input vector $x = (i, d, y_{it-1}, \ldots)$ where i is the box id, d is the district id, and $\{y_{it-1}, \ldots\}$ is 1 years of crime lags for all crime types used to predict All Crimes PredPol boxes. The multilayer perceptron neural network is a kind of so-called deep neural network, which is a universal function approximator that thrives in large-scale data settings. I implement and train the neural network model using the open-source Keras/Tensorflow python libraries.[23] The neural network has two fully-connected hidden layers, followed by rectified linear and sigmoidal activation functions, respectively.[24] For variables that take discrete values (discrete variables), I use an embedding function which maps discrete variables to continuous features. For instance, PredPol includes a time-invariant box parameter in their crime risk model; I model this using an embedding space for both the district and box id discrete variables, e.g. I map the discrete d and box i index to learned vector representations. There are more non-PredPol boxes than PredPol boxes; to address this imbalance, I use the class count weight to weigh the loss function (Keras/Tensorflow feature).

The neural network outputs a likelihood of PredPol box probability $\hat{\lambda}_{it}$ between 0 and 1

[23]http://tensorflow.org. I used version 2.4.1 with GPU support.

[24]I explored the number and size of the MLP hidden layers, as well as the stochastic gradient descent and adaptive moment estimation (Adam) learning algorithms.

for every observation x:

$$AC^{pred} = \begin{cases} 0 & \text{if } \hat{\lambda}_{it}(x) < 0.5 = \tau \\ 1 & \text{if } \hat{\lambda}_{it}(x) \geq 0.5 = \tau \end{cases} \tag{10}$$

I use a two-year sample of ever-PredPol boxes from 3/1/2018-3/1/2020, with 1 year of crime lags for the crime types used to predict All Crimes PredPol boxes. I randomly split the sample into a training set to train the neural network (60% of sample) and a test set to test model performance out-of-sample (40% of sample). Table 6 shows the performance of the neural network in prediction accuracy in the test set, defined as the percent of boxes for which PredPol box treatment status is correctly predicted. The best model is defined as the model that has the best overall prediction accuracy (on the test set) that also equalizes prediction accuracy for PredPol boxes and non-PredPol boxes. The best performance in the test set achieves around 92.14% prediction accuracy overall. Boehnke and Bonaldi (2019) also use machine learning for their high first stage prediction, achieving 0.979 accuracy in their validation set. The size of the data and the memory it requires to train the model limited how extensive the training and investigation of the MLP predictor could be. In the future, it may be possible to improve on the prediction accuracy.

6.2 Second stage: Regression discontinuity design using the synthetic running variable

In the second stage, I use the synthetic running variable $\hat{\lambda}_{it}$ as the running variable in a sharp regression discontinuity design framework conditional on treatment status, estimated in the test set data:

$$Y_{it} = \alpha + \beta AC_{it} + P(\hat{\lambda}_{it}) + \varepsilon_{it} \tag{11}$$

where Y_{it} is the crime incidence in box i at time t and AC_{it} is the indicator for whether box i is an All Crimes PredPol box at time t, and $P(\hat{\lambda}_{it})$ is a polynomial function of the estimate of the continuous risk score, or synthetic running score, $\hat{\lambda}_{it}$. β is the local average treatment effect at the margin of All Crimes PredPol box treatment conditioning on treatment status. Following Boehnke and Bonaldi (2019), I drop misclassified boxes to guarantee the discontinuity in probability of PredPol box treatment at τ, the threshold of $\hat{\lambda}_{it}$. I further restrict the sample to the period before 11/20/2019 when All Crimes PredPol boxes are active during the day.[25] I account for within-box correlation of errors over time with clustered standard errors. Following Boehnke and Bonaldi (2019), I use the bias-corrected RD estimator of Calonico et al. (2014), to "perform inference that is robust to the choice of bandwidth for the estimation of the local polynomials near the threshold" (Boehnke and Bonaldi, 2019).[26]

6.3 Results

Table 7 shows the effects of algorithm-induced police presence (Active All Crimes PredPol Box coefficient) on serious property and violent crime incidents, including aggravated assault, burglary, robbery and vehicle theft. Column (1) shows the results from my main empirical strategy presented in Section 4, β from equation 2. Column (2) applies the synthetic regression discontinuity (SRDD) design framework of Boehnke and Bonaldi (2019) to further examine the effect of predictive policing box presence on crime as a robustness check. Column (2) finds a reduction of 4.954 serious property and violent crime incidents per 1000 boxes, which is a 63.7% reduction relative to the PredPol box mean of what to expect during the day shift, 7.773 crimes per 1000 boxes. The two frameworks isolate different sources of quasi-experimental variation to estimate treatment effects. Moreover, SRDD estimates the local average treatment effect of algorithm-induced police presence at predictive policing boxes around the threshold of treatment. I find that the estimates from both frameworks have the same sign, providing a robustness for the main empirical strategy and compelling

[25] For this period, All Crimes PredPol boxes are active during the day and $AC_{it} = Active\_AC_{it}$.

[26] This draft does not yet account for first-stage variation in second stage standard errors.

support for my conclusions on the effects of algorithm-induced police presence.

7 Conclusion

This paper investigates the effects of algorithm-induced police presence on crime incidents and racial disparities in arrests. As predictive policing algorithms become ubiquitous, it is important to understand the local impacts of algorithm-induced police presence. I isolate quasi-experimental variation using two natural experiments to study the causal impacts of algorithm-induced police presence. First, I validate that law enforcement is responding to predictive policing boxes, finding an increase in shots fired calls for service in predictive policing boxes due to algorithm-induced police presence. Second, my findings indicate that algorithm-induced police presence deters serious violent and property crime; I also find suggestive evidence that algorithm-induced police presence increases the number of traffic incidents, a lower-level offense where law enforcement has discretion. Third, there is also evidence that algorithm-induced police presence has disproportionate racial impacts on arrests for serious violent crimes in PredPol boxes and surrounding boxes, and arrests in traffic incidents in PredPol boxes.

These results imply that the impacts of algorithm-induced police presence on crime differ by crime type, which has implications for concerns about feedback loops in algorithms. Of the violation code types used to generate predictive policing boxes, I find that algorithm-induced police presence increases shots fired calls for service incidents, and decreases serious property and violent crimes. In the jurisdiction that I study, traffic incidents are not used to generate predictive policing boxes. However, I find suggestive evidence that algorithm-induced police presence increases reported traffic incidents, underlining that lower-level offenses where police have discretion are of particular concern for this algorithmic feedback loop.

Moreover, while there is evidence that algorithm-induced police presence deters crime for certain crime types, there are important equity implications of using predictive policing to target areas, as there is also evidence that Black arrestees are disproportionately arrested for certain crimes as a result of algorithm-induced police presence. In the future, I aim to use similar data that I have collected from other US cities to extend this analysis and speak to the external validity of this analysis. Moreover, the empirical strategies that I develop and use can be applied to data from other cities to further study impacts of algorithmic policing, and to other contexts where algorithmic decision systems are used to measure policy-relevant treatment effects.

References

Becker, G. S. (1957). *The Economics of Discrimination: An Economic View of Racial Discrimination*. University of Chicago.

Blanes i Vidal, J. and Mastrobuoni, G. (2018). Police patrols and crime.

Blattman, C., Green, D. P., Ortega, D., and Tobón, S. (2021). Place-based interventions at scale: The direct and spillover effects of policing and city services on crime. *Journal of the European Economic Association*, 19(4):2022–2051.

Boehnke, J. and Bonaldi, P. (2019). Synthetic regression discontinuity: Estimating treatment effects using machine learning.

Braga, A. A., Turchan, B. S., Papachristos, A. V., and Hureau, D. M. (2019). Hot spots policing and crime reduction: an update of an ongoing systematic review and meta-analysis. *Journal of experimental criminology*, 15(3):289–311.

Brantingham, P. J., Valasik, M., and Mohler, G. O. (2018). Does predictive policing lead to biased arrests? results from a randomized controlled trial. *Statistics and Public Policy*, 5(1):1–6.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Chalfin, A., Hansen, B., Weisburst, E. K., Williams, M. C., et al. (2020). Police force size and civilian race. Technical report, National Bureau of Economic Research.

Chalfin, A. and McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, 55(1):5–48.

Di Tella, R. and Schargrodsky, E. (2004). Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, 94(1):115–133.

Draca, M., Machin, S., and Witt, R. (2011). Panic on the streets of london: Police, crime, and the july 2005 terror attacks. *American Economic Review*, 101(5):2157–81.

Durlauf, S. N. and Nagin, D. S. (2011). Imprisonment and crime: Can both be reduced? *Criminology & Public Policy*, 10(1):13–54.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. (2017). Human decisions and machine predictions. *The quarterly journal of economics*, 133(1):237–293.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Sunstein, C. R. (2018). Discrimination in the age of algorithms. *Journal of Legal Analysis*, 10.

Klick, J. and Tabarrok, A. (2005). Using terror alert levels to estimate the effect of police on crime. *The Journal of Law and Economics*, 48(1):267–279.

Lum, K. and Isaac, W. (2016). To predict and serve? *Significance*, 13(5):14–19.

Mastrobuoni, G. (2020). Crime is terribly revealing: Information technology and police productivity. *The Review of Economic Studies*, 87(6):2727–2753.

Mello, S. (2019). More cops, less crime. *Journal of Public Economics*, 172:174–200.

Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., and Brantingham, P. J. (2015). Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, 110(512):1399–1411.

O'Neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Ratcliffe, J. H., Taylor, R. B., Askey, A. P., Thomas, K., Grasso, J., Bethel, K. J., Fisher, R., and Koehnlein, J. (2020). The philadelphia predictive policing experiment. *Journal of Experimental Criminology*, pages 1–27.

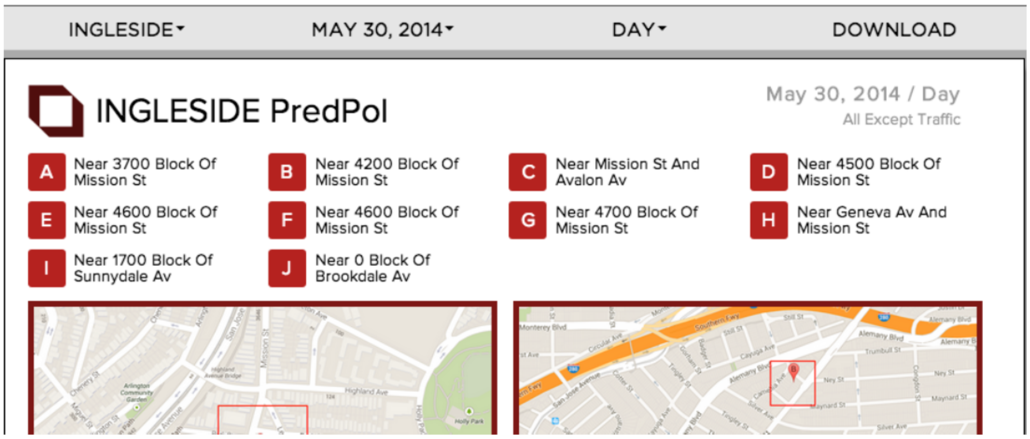
Stevenson, M. and Doleac, J. (2019). Algorithmic risk assessment tools in the hands of humans.

Weisburd, D. and Telep, C. W. (2014). Hot spots policing: What we know and what we need to know. *Journal of Contemporary Criminal Justice*, 30(2):200–220.

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

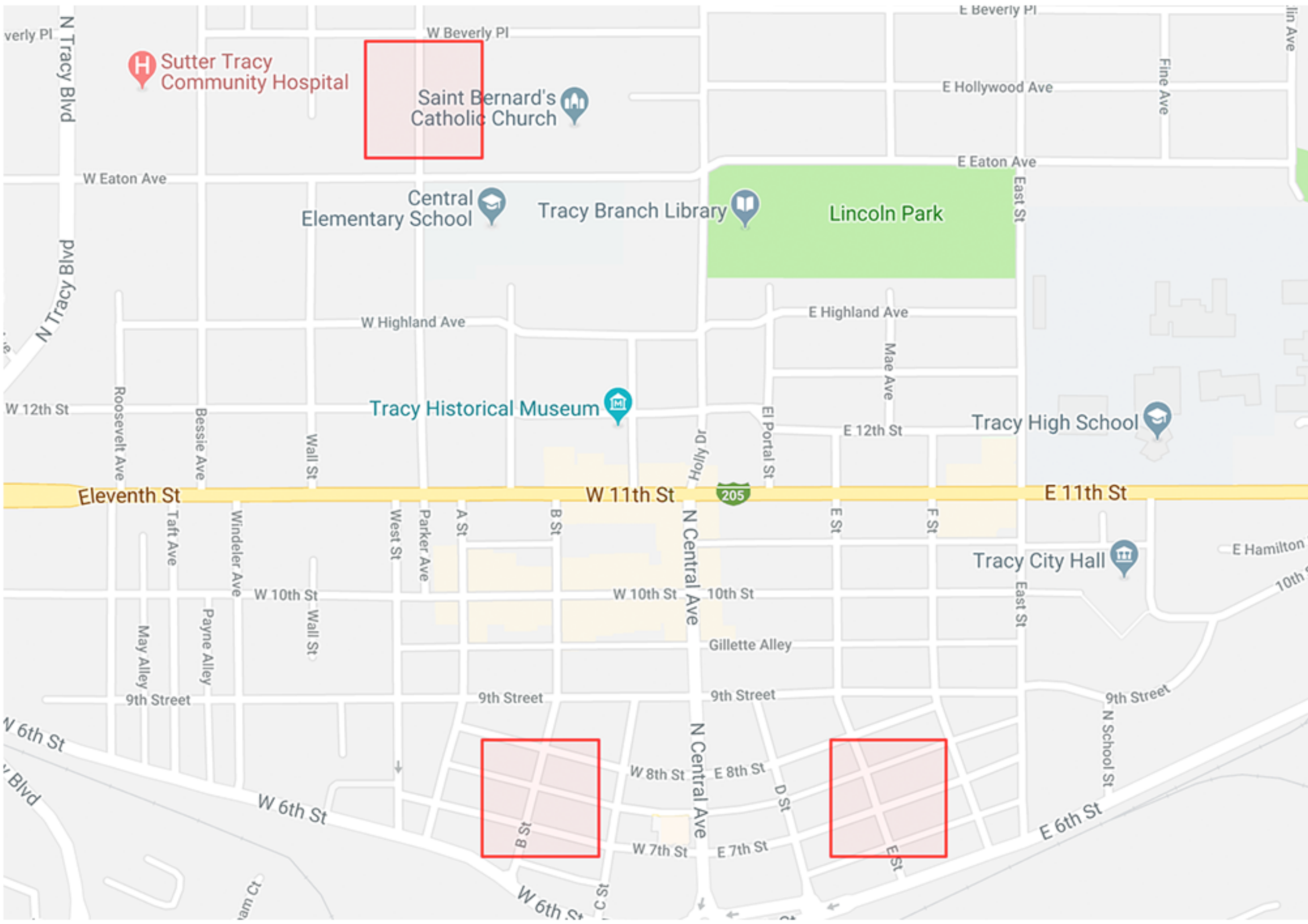
Figures and Tables

Figure 1: Example of a PredPol patrol report from PredPol guide



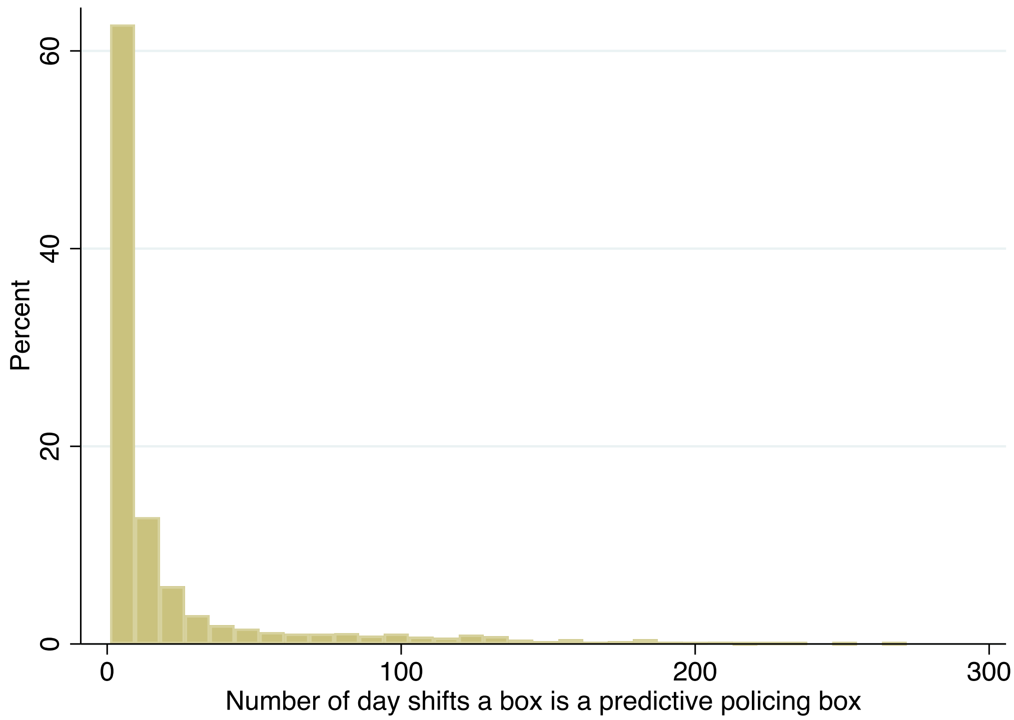
Notes: This is an example of a PredPol patrol report from a PredPol guide to illustrate what law enforcement see, and does not necessarily come from the jurisdiction that I study in this paper. PredPol boxes are shown on maps and also identified using the approximate intersection. The patrol report is for May 30, 2014 for the Day shift, for all except traffic crimes.

Figure 2: Map of PredPol predictive policing boxes from PredPol website



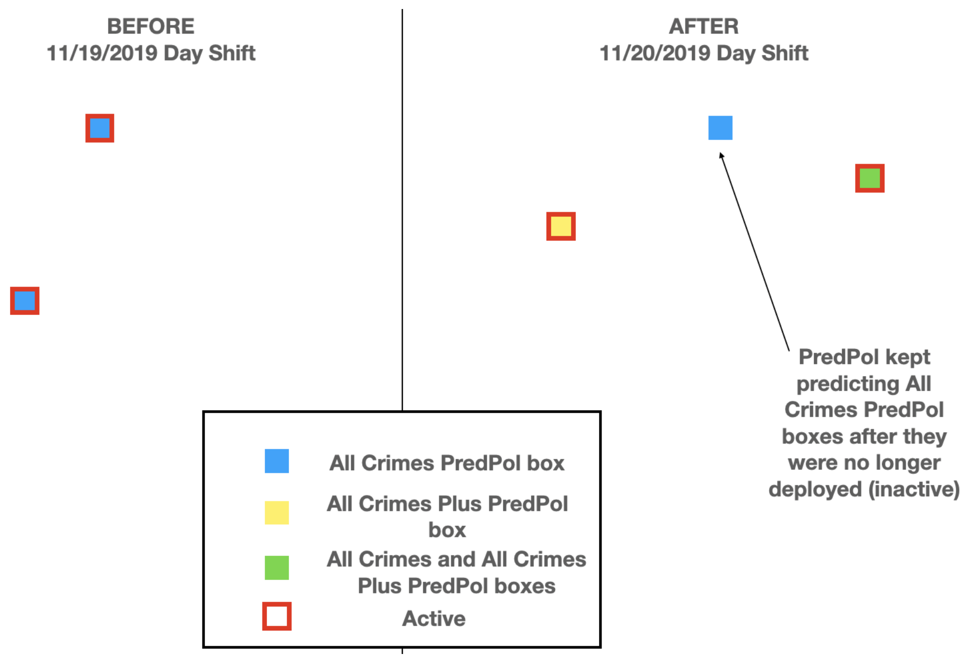
Notes: This is an example of a map of PredPol boxes from the PredPol website. PredPol boxes are the red boxes. This figure illustrates the size of PredPol boxes which are 500 feet by 500 feet and span 1-3 blocks. The map does not necessarily come from the jurisdiction that I study in this paper.

Figure 3: Distribution of number of day shifts a box is a predictive policing box over the quasi-random experiment window (5/20/2019-3/1/2020)



Notes: This figure plots the number of day shifts a box is a All Crimes predictive policing box over the quasi-random experiment window (5/20/2019-3/1/2020) for the 1,924 boxes that will be an All Crimes predictive policing boxes over the quasi-random experiment window (5/20/2019-3/1/2020).

Figure 4: Illustration of active predictive policing box quasi-experiment research design



Notes: Prior to 11/20/2019, law enforcement receives the All Crimes PredPol boxes in blue, and law enforcement is instructed to patrol in these boxes. After 11/20/2019, PredPol kept predicting All Crimes PredPol boxes even when they were delivered to law enforcement, and they serve as a control group in my research design. After 11/20/2019, law enforcement receives the All Crimes Plus PredPol boxes in yellow, and law enforcement is instructed to patrol in these boxes. There is also overlap between the All Crimes and All Crimes Plus PredPol boxes after the change since there is overlap in the crime types used in prediction. My research design compares the outcomes at the the All Crimes PredPol boxes before the change with the outcomes at the All Crimes Plus PredPol boxes after the change (which are not delivered and therefore law enforcement is not instructed to patrol there), accounting for the All Crimes Plus PredPol boxes that are delivered to law enforcement after the change (where they are instructed to patrol) and box fixed effects.

Table 1: The effect of algorithm-induced police presence on shots fired calls for service (police activity)

	(1)	(2)	(3)
Active All Crimes PredPol Box	0.513**	0.564**	0.578**
	(0.258)	(0.260)	(0.261)
Outcome Mean	0.143	0.143	0.143
PredPol Box Outcome Mean	0.486	0.486	0.486
Box ID Fixed Effects	Yes	Yes	Yes
Underlying crime risk	Lags	Lags	Lags
District-Time	No	Trends	Fixed Effects
Clusters	8224	8224	8224
Observations	2352064	2352064	2352064

Notes: This table presents estimates of β from equation 2. Sample of all box-shifts of boxes that are ever All Crimes and All Crimes PredPol boxes over a three year period. Regressions control for crime lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Lags include 7 day shift lags and 12 month lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Standard errors are in parentheses and are clustered at box level. Stars signify: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. Estimates and outcome means are multiplied by 1000.

Table 2: The effect of algorithm-induced police presence on serious property and violent crime incidents

	Box			Expanded box
	(1)	(2)	(3)	(4)
Active All Crimes PredPol Box	-3.077**	-2.943**	-2.909**	-4.844**
	(1.474)	(1.478)	(1.478)	(2.244)
Outcome Mean	1.685	1.685	1.685	8.847
PredPol Box Outcome Mean	9.348	9.348	9.348	25.325
Box ID Fixed Effects	Yes	Yes	Yes	Yes
Underlying crime risk	Lags	Lags	Lags	Lags
District-Time	No	Trends	Fixed Effects	Fixed Effects
Clusters	8224	8224	8224	8224
Observations	2352064	2352064	2352064	2352064

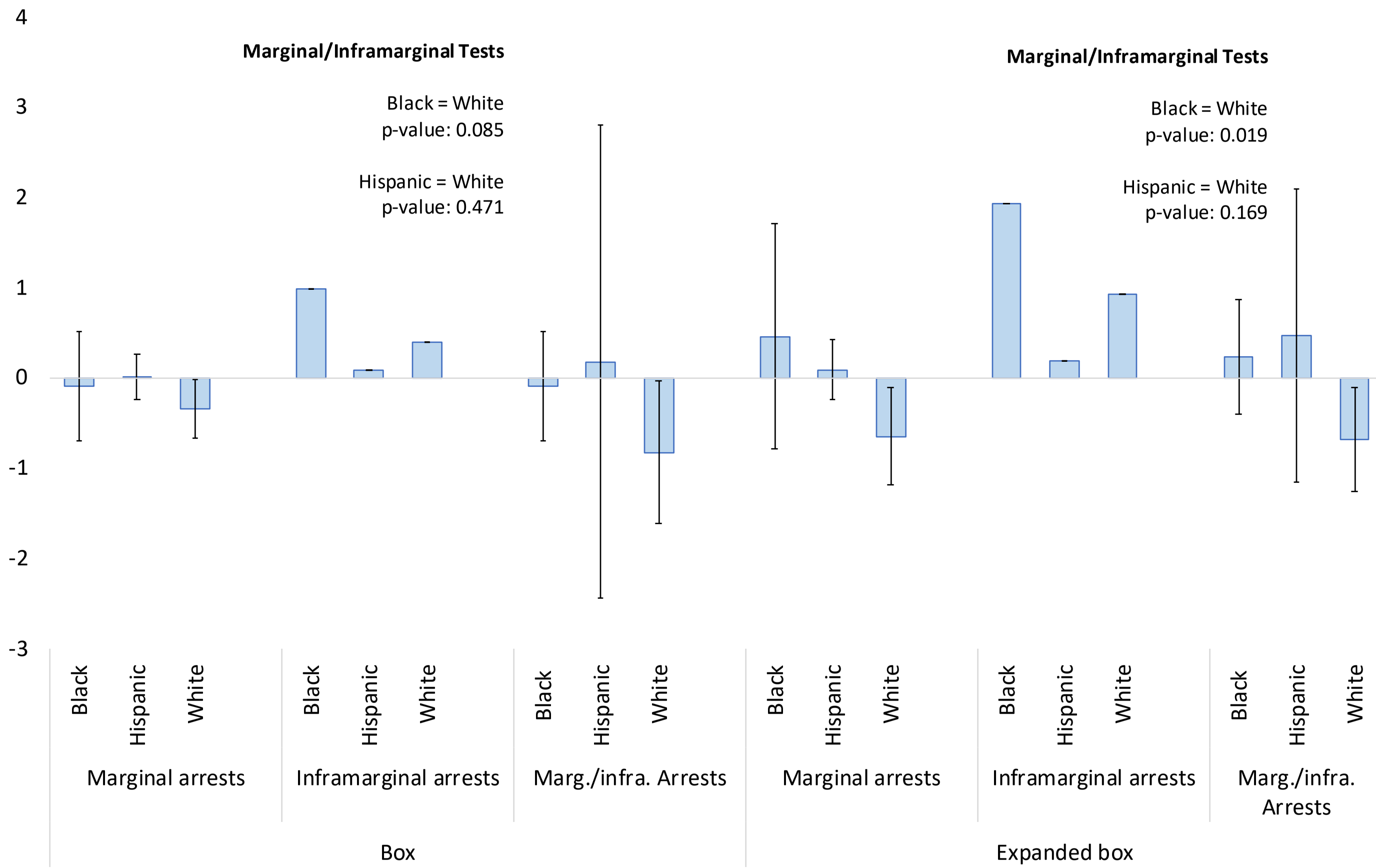
Notes: This table presents estimates of β from equation 2. Serious property and violent crime incidents include aggravated assault, burglary, robbery, and motor vehicle theft. Sample of all box-shifts of boxes that are ever All Crimes and All Crimes Plus PredPol boxes over a three year period. Regressions control for crime lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Lags include 7 day shift lags and 12 month lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Standard errors are in parentheses and are clustered at box level. Stars signify: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. Estimates and outcome means are multiplied by 1000.

Table 3: The effect of algorithm-induced police presence on traffic incidents

	(1)
Active All Crimes PredPol Box	0.779*
	(0.428)
Outcome Mean	0.393
PredPol Box Outcome Mean	1.675
Box ID Fixed Effects	Yes
Underlying crime risk	Lags
District-Time	Fixed Effects
Clusters	8224
Observations	2352064

Notes: This table presents estimates of β from equation 2. Sample of all box-shifts of boxes that are ever All Crimes and All Crimes Plus PredPol boxes over a three year period. Regressions control for crime lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Lags include 7 day shift lags and 12 month lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Standard errors are in parentheses and are clustered at box level. Stars signify: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. Estimates and outcome means are multiplied by 1000.

Figure 5: Test for disproportionate racial impacts of algorithm-induced police presence on arrests of serious violent crime



Notes: The left panel shows estimates for predictive policing boxes, and the right panel shows estimates for predictive policing boxes and surrounding boxes (expanded box outcome). The left set of bars per panel plots the estimates of the effect of algorithm-induced police presence by race (β_w, β_b, β_h from equation 4); the center set of bars per panel plots the number of inframarginal arrests by race ($y_{cf,w}$, $y_{cf,b}$, and $y_{cf,h}$ from equation 6); the right set of bars per panel plots the effect weighted by the number of inframarginal arrests by race. P-values are from two-sided tests testing equation 7.

Table 4: Test for disproportionate racial impacts of algorithm-induced police presence on arrests of serious violent crime

	Box (1)	Expanded box (2)
White effect β_w	-0.337**	-0.638**
	(0.166)	(0.277)
Black effect β_b	-0.084	0.470
	(0.307)	(0.635)
Hispanic effect β_h	0.018	0.098
	(0.127)	(0.169)
White inframarginal arrests $y_{cf,w}$	0.412	0.940
Black inframarginal arrests $y_{cf,b}$	0.990	1.945
Hispanic inframarginal arrests $y_{cf,h}$	0.095	0.204
White effect/ infra.	-0.818	-0.679
Black effect/ infra.	-0.084	0.242
Hispanic effect/ infra.	0.189	0.480
P-value: Black effect/ infra.= White effect/ infra.	0.085	0.019
P-value: Hispanic effect/ infra. = White effect/ infra.	0.471	0.169
Box ID Fixed Effects	Yes	Yes
Underlying crime risk	Lags	Lags
District-Time	Fixed Effects	Fixed Effects
Clusters	8224	8224
Observations	7056192	7056192

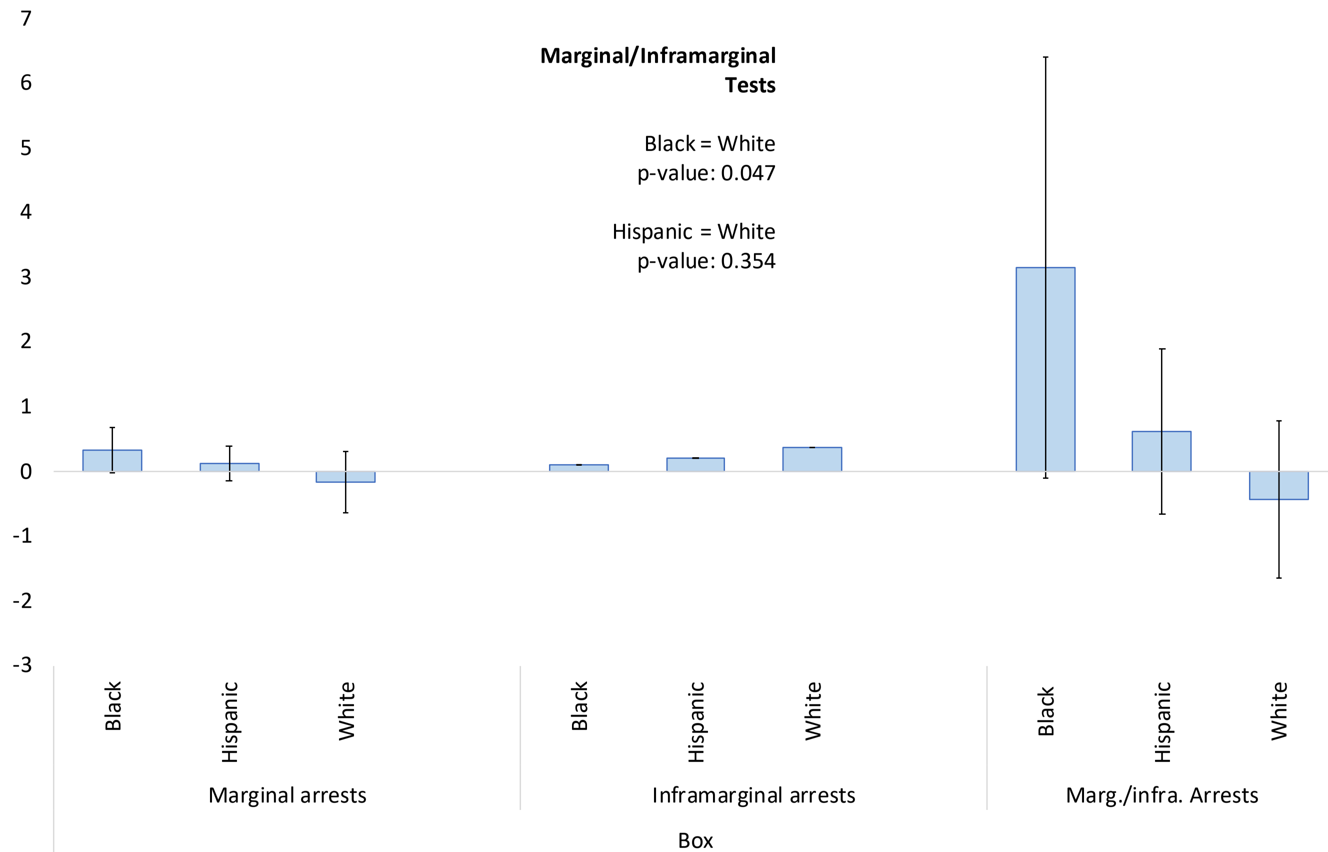
Notes: This table presents estimates of β_w, β_b, β_h, $y_{cf,w}$, $y_{cf,b}$, and $y_{cf,h}$ from equation 4 and equation 6. P-values are from two-sided tests testing 7. Serious violent crime includes aggravated assault and robbery. Sample of all box-shifts of boxes that are ever All Crimes and All Crimes Plus PredPol boxes over a three year period. Regressions control for crime lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Lags include 7 day shift lags and 12 month lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Standard errors are in parentheses and are clustered at box level. Stars signify: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. Estimates and outcome means are multiplied by 1000.

Table 5: Test for disproportionate racial impacts of algorithm-induced police presence on arrests of serious property crime

	(1)
White effect β_w	-0.550*
	(0.284)
Black effect β_b	-0.888*
	(0.506)
Hispanic effect β_h	-0.208
	(0.992)
White inframarginal arrests $y_{cf,w}$	0.550
Black inframarginal arrests $y_{cf,b}$	1.265
Hispanic inframarginal arrests $y_{cf,h}$	0.359
White effect/ infra.	-1.000
Black effect/ infra.	-0.702
Hispanic effect/ infra.	-0.579
P-value: Black effect/ infra.= White effect/ infra.	0.639
P-value: Hispanic effect/ infra. = White effect/ infra.	0.863
Box ID Fixed Effects	Yes
Underlying crime risk	Lags
District-Time	Fixed Effects
Clusters	8224
Observations	7056192

Notes: This table presents estimates of β_w, β_b, β_h, $y_{cf,w}$, $y_{cf,b}$, and $y_{cf,h}$ from equation 4 and equation 6. P-values are from two-sided tests testing 7. Serious property crimes include residential burglary, commercial burglary, auto burglary, and motor vehicle theft. Sample of all box-shifts of boxes that are ever All Crimes and All Crimes Plus PredPol boxes over a three year period. Regressions control for crime lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Lags include 7 day shift lags and 12 month lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Standard errors are in parentheses and are clustered at box level. Stars signify: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. Estimates and outcome means are multiplied by 1000.

Figure 6: Test for disproportionate racial impacts of algorithm-induced police presence on arrests in incidents of traffic nature



Notes: The left set of bars plots the estimates of the effect of algorithm-induced police presence by race (β_w, β_b, β_h from equation 4); the center set of bars plots the number of inframarginal arrests by race ($y_{cf,w}$, $y_{cf,b}$, and $y_{cf,h}$ from equation 6); the right set of bars plots the effect weighted by the number of inframarginal arrests by race. P-values are from two-sided tests testing equation 7.

Table 6: Multilayer perceptron neural network predictive accuracy of underlying crime risk running score in test set

	Multilayer perceptron predictive accuracy
Non-PredPol boxes	92.19%
PredPol boxes	92.11%
Overall	92.14%

Notes: I use a multilayer perceptron (MLP) neural network to predict whether a box i at time t is a PredPol box using input vector $x = (i, d, y_{it-1}, \dots)$ where i if the box id, d is the district id, and $\{y_{it-1}, \dots\}$ is 1 years of crime lags for all crime types used to predict All Crimes PredPol boxes. The neural network has two fully-connected hidden layers, followed by rectified linear and sigmoidal activation functions, respectively.

Table 7: Effect of algorithm-induced police presence on serious index crime incidents estimated using synthetic regression discontinuity design

	Active PredPol Box Quasi-Experiment	Synthetic RDD
	(1)	(2)
Active All Crimes PredPol Box	-2.909**	-4.954***
	(1.478)	(1.337)
Conventional p-value		0.000
Robust p-value		0.000
Outcome Mean	1.685	1.769
PredPol Box Outcome Mean	9.348	7.773
Box ID Fixed Effects	Yes	No
Underlying crime risk	Lags	SRDD
District-Time	Fixed Effects	No
Clusters	8224	
Observations	2352064	1344148

Notes: Serious index crime incidents include aggravated assault, burglary, robbery, and vehicle theft. Standard errors are in parentheses and are clustered at box level. Standard errors for synthetic regression discontinuity design models (Boehnke and Bonaldi, 2019) in Column (2) does not account for variation in the first stage prediction outcome. Column (1) controls for 7 day shift lags and 12 month lags summing the crimes included in prediction for All Crimes and All Crimes Plus PredPol boxes. Estimates and outcome means are multiplied by 1000. Stars signify: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.