Race, Criminal History, and Prosecutor Case Selection:

Evidence from a Southern U.S. Jurisdiction

J.J. NADDEO[*][†][‡]

November 3, 2022

Abstract

State prosecutors are powerful actors who handle the vast majority of criminal cases in the United States. Furthermore, they are located at the core of the criminal justice pipeline, which runs from arrest to sentencing, and produces large persistent racial disparities. Given their centrality and influence, it is important to understand how prosecutors exacerbate or mitigate these disparities. This paper uses administrative data developed in collaboration with a prosecutor's office located in a medium-sized jurisdiction in the southern United States. Using both descriptive and causal methods, I find that prosecutors are complex actors that are aware of upstream biases that taint signals they receive and act to *reduce* racial disparities. Specifically, after controlling for a rich set of covariates, I find that Black individuals receive shorter sentences and are more likely to have their charges dismissed relative to similarly situated White individuals. To learn more about what is happening at the margin, I leverage quasi-random assignment of cases to prosecutors and a simple—yet novel—model to show that prosecutors discount how prior convictions map into punishment for Black individuals relative to White individuals. My results suggest policies aimed at removing prosecutorial discretion or "blinding" prosecutors from knowing the race of an individual may have unintended consequences.

[*]Massive Data Institute, McCourt School of Public Policy, Georgetown University
[†]Institute for Technology Law & Policy, Georgetown University Law Center
[‡]Justice Innovation Lab

1 Introduction

The U.S. criminal legal system is marked by large and persistent racial disparities. By age 23, nearly half of all Black men have been arrested, compared with only a third of White men (Brame et al., 2014). Similarly, over 25% of all Black people have been convicted of a felony, compared with 6% of all non-Black individuals (Shannon et al., 2010).[1]

In theory, these disparities may arise from differences in actual underlying criminal activity, potentially rooted in generational deficits caused by historical discrimination, or driven by a constellation of contemporary upstream biases (e.g., discrimination in labor markets, access to higher education, etc.). But a long literature suggests these persistent racial disparities may stem more directly from contemporary decisions of actors in the criminal legal system.[2] Although much of this research has focused on the roles of judges and police officers, a growing number of papers have examined prosecutors.[3]

This paper focuses on state prosecutors, who handle the vast majority of criminal cases in the U.S. and are widely recognized as the most influential actors in the system.[4] Moreover, prosecutors are embedded in the core of the criminal legal pipeline from arrest to sentencing and are aware of racial bias in upstream processes, such as policing (Dunlea, 2021; Shaffer, 2022). Given their power, centrality, and capacity to exercise discretion, understanding how race affects prosecutorial decision-making directly and indirectly through upstream signals

[1]Recent data from the FBI Uniform Crime Report corroborate these patterns (FBI Uniform Crime Report, 2019). They illustrate 26.6% of individuals arrested are Black, whereas only 14% of the total population is Black. Although this rate has declined over recent years, the imprisonment rate of Black adults is still five times larger than that of White adults (Carson, 2020).

[2]See Doleac (2022) for a recent review of the causal empirical evidence of bias from different actors and decision points.

[3]See (Rehavi and Starr, 2014; Tuttle, 2021; Sloan, 2022; Shaffer, 2022) for more recent papers and Wu (2016) for a meta-analysis focused on papers published before 2012.

[4]See Bibas (2009); Sklansky (2016); Pfaff (2017); American Civil Liberties Union (2022).

is crucial.

Recent reform efforts have focused on mitigating potential prosecutorial bias by removing race as a salient factor in prosecutors' decision-making process (e.g., "color-blinding" by scrubbing race from intake forms).[5] Such efforts, however, could have unintended consequences, such as increased racial disparities. A recent example comes from attempts to "ban-the-box", where removing information about applicants' criminal histories *increased* racial disparities in employment and recidivism (Agan and Starr, 2018; Doleac and Hansen, 2020; Sherrard, 2020).

This paper leverages a novel high-quality administrative dataset from a mid-sized jurisdiction in the southern U.S. to study the impact of prosecutorial discretion on racial disparities. These records track cases from arrest to disposition, allowing for a fuller picture of how prosecutors use race, given that they are only one part of the criminal legal pipeline. Namely, I am able to decompose observed raw disparities into components driven by bias from prosecutors and a residual systemic bias that is driven by upstream forces. Informed by evidence from this detailed descriptive analysis, I use a policy experiment and quasi-random variation in case assignment to prosecutors to obtain causal estimates of how prosecutors make indictment decisions. I further decompose these causal estimates using a novel theoretical model that allows me to identify the causal effect of additional convictions on the probability of conviction accounting for arrest patterns. Overall, I find contemporary prosecutors are unlikely to drive observed racial disparities and, in fact, act as a countervailing force by reducing the weight placed on criminal histories contaminated by upstream biases.

The insights of this paper come from the quality and construction of the underlying dataset. The data come from one particular jurisdiction, where administrative staff and prosecutors helped me create a data pipeline. The process, guided by the agency itself,

---

[5]On September 29, 2022, California Governor Gavin Newsom approved Race Blind Charging bill AB 2778, which unanimously passed both state legislative bodies. The bill mandates race-blind charging in all district attorneys' offices by 2025. See Sah et al. (2015); Chohlas-Wood et al. (2021) for academic papers on the theory and implementation of the policy.

minimized misrepresentation of the data as well as coding errors. These processes allow me to reduce measurement error and provided contextual details unavailable to many other researchers.[6] These records capture an individual's demographic features, such as their criminal history, residential address, age, gender, and race (as identified by the office). Importantly, I am also able to observe the exact charges booked by the arresting officer, as well as the outcome of each charge and other charge characteristics, including the number of victims and co-defendants. My data include any crime that carries the penalty of incarceration in a state prison, which represents the vast majority of incarcerated people, because federal prisons only represent around 5% of the total number of incarcerated individuals at any given time over the past 40 years.[7]

Representativeness of data is crucial for making inferences that apply elsewhere. The jurisdiction from which the data originates boasts two attractive features that reinforce external validity. First, prosecutors—as elected officials—are bound to represent the electorate they serve. As a consequence, their constituents' preferences likely dictate prosecutorial behavior to some extent. Such sensitivity increases the likelihood that patterns in prosecutorial behavior are heterogeneous across both time and space. The jurisdiction of interest contains voters with relatively moderate preferences; as such, its Republican chief prosecutor does not exhibit radical tendencies in sentencing determination. Indeed, most elected state district attorneys run on non-progressive platforms and are affiliated with the Republican party (Stashko and Garro, 2021).[8] Second, population density plays a significant role in how intuitions develop. For example, large densely populated areas likely have prosecutor's offices with larger budgets and more line staff and that experience a different distribution of crime.

---

[6]For instance, data on cases with a guilty adjudication may lack details on sentencing. Code was written to identify hundreds of such instances. The staff then sourced the hand-written sentencing sheets and adjusted the data.

[7]According to data compiled by the Prison Policy Initiative, the share of incarcerated people due to state and local prosecution ranged from 91% to 96% in 1983-2020. Data accessed here.

[8]No clear consensus on what defines a progressive jurisdiction exists. See definitions in Stashko and Garro (2021); Agan et al. (2022); Hogan (2022); Mitchell et al. (2022). Despite a lack of consensus, each of the previously cited definitions would classify the jurisdiction of interest as non-progressive.

The existing literature focuses predominantly on the most densely populated urban areas, such as New York City, Philadelphia, Miami, and Boston. Conversely, the location where data come from is a medium-sized urban area with approximately 600,000 people.[9]

Consistent with the literature to date, I find Black men's sentences are 15% longer than the sentences of similarly aged White men (Rehavi and Starr, 2014). A rich set of observables enables me to gauge the extent to which prosecutors' decision-making influences this pattern. Specifically, I examine the correlation between race and length of sentence conditional on the number of prior convictions, exact arrest charge, number of charges brought by police, number of victims, number of co-defendants, and defense counsel characteristics. After conditioning for these characteristics, I find Black men receive sentences that are 10% *shorter* on average than sentences for similarly situated White defendants. This result suggests prosecutors *reduce* the racial difference in sentence lengths on average.

To further decompose what drives the dramatic differences between the conditional and unconditional racial disparity in incarceration, I utilize the framework set forth in Bohren et al. (2022). This strategy offers a potential means of disentangling the influence of direct and systematic discrimination. Direct discrimination refers to the effect a group characteristic (*i.e.*, race in this context) has on outcomes. By contrast, systemic discrimination represents a decision-maker's bias stemming from how they make use of variables that are biased by upstream forces. For example, prosecutors may rely heavily on an individual's criminal history, which can be a product of layers of current and historical racially biased decisions. In the context of my results, the unconditional racial disparity of 15% represents a measure of total disparity, whereas the best estimate of direct bias from prosecutors is -10%. Therefore, the 25% difference represents an estimate of systemic bias, or bias driven by differences across legally permissible observables.

To empirically measure which observables drive this upstream bias, I use the decomposi-

---

[9]According to the most recent survey of state prosecutors in Perry and Banks (2011), only 43 of the 2,330 (1.8%) state prosecutor's offices in the U.S. served populations greater than 1 million.

tion method from Gelbach (2016). My results indicate the majority of bias (64%) comes from the number of prior convictions an individual enters the system with. This finding implies the distribution of prior convictions explains a large portion of observed racial disparities in incarceration lengths. From discussions with prosecutors in this office, an individual's rap sheet (i.e., criminal record) is indeed one of the most salient pieces of information that influence a prosecutor's decisions. Additionally, I find Black individuals are much more likely to have more extensive criminal records. Therefore, understanding how race influences the prosecutor's mapping of prior convictions into decisions is important. I find that *ceteris paribus* (with respect to the aforementioned set of controls), additional prior convictions increase the average incarceration more for White defendants than for similarly situated Black defendants. Specifically, an increase from no prior convictions to three prior convictions increases the average carceral sentence of a White individual by 45% compared with an 18% increase for a similarly situated Black individual.[10,11]

Prosecutors enjoy almost unchecked power when deciding whether to prosecute charges. Therefore, I investigate the influence of such power by comparing dismissal rates by race for individuals with similar observable characteristics. I find that, on average, Black men are approximately 3 percentage points (10%) more likely to have full or partial case dismissal than similarly situated White men.

However, OLS estimates come from comparisons of individuals with identical observables (such as, arrest charges, criminal history, etc.), however it is likely that at the margin, Black and White individuals differ across these observables. To get better leverage on what cases are being pursued at the margin, I turn to a pilot screening program employed in the jurisdiction where the data originate. An individual—aptly referred to as a screener—reviews

---

[10]This result is not driven by the composition of prior convictions, because I find no racial differences in the percentage of priors that are felonies across the distribution of prior convictions (e.g., on average, both Black and White individuals with four prior convictions have two felonies and two misdemeanors.

[11]This result is remarkably similar to the results in Shaffer (2022), in which prior convictions increased the probability of imprisonment more for White individuals than for Black individuals. Shaffer (2022) also provides survey evidence that these effects are concentrated in younger prosecutor cohorts that believe are more aware of racial bias in the criminal justice system.

cases and recommends cases to the prosecutor. The screener is given a clear mandate to remove low-level "junk" cases that are unlikely to result in conviction. In addition, I was able to provide the screener with a survey in which they could express their reasoning behind their decisions. From the responses, the screener was clearly focused primarily on removing cases that they believed would not lead to successful cases and/or were not in the interest of justice.[12] For these reasons, I assume the screener is likely to use race-neutral thresholds when making decisions. To test for racial bias in historical decisions, I compare historical prosecution rates with the prosecution rates of the screener. If, for example, prosecutors set race-neutral dismissal requirements historically, I would expect the change in dismissal rates caused by the screener to exhibit no racial heterogeneity.[13] Instead, I find substantial racial heterogeneity in the effect of screening on prosecution, with the prosecution rate decreasing for Black individuals and remaining unchanged for White individuals. This simple pre/post analysis may be biased if the portfolio of charges were changing over time. So, to confirm my results, I construct counterfactual control groups from non-eligible arresting agencies in a difference-in-differences framework as well as a synthetic control framework, where I find almost identical estimates. These results suggest that even though prosecutors dismiss charges involving Black individuals at a higher rate on average, this rate is insufficient to equalize the marginal case brought against Black individuals.

In addition to the screening unit, I am also able to exploit variation driven by quasi-random assignment of cases to prosecutors to learn about prosecutorial decision-making at the margin. I utilize this variation in an examiner design ("judge IV") framework to recover causal estimates of how an additional conviction impacts the probability an individual is re-arrested and re-convicted (conditional on arrest). Using either future re-arrest or re-

---

[12]For example, the majority of reasoning generally involved mention of pre-textual stops, small amounts of drugs, and gun possession. Subsequent interviews with the screener also re-enforced these qualitative observations. Note a pretextual stop is where law enforcement initiates contact for something like a broken taillight but then finds drugs and makes an arrest.

[13]If the screener is more (less) lenient than the status quo, I should observe parallel changes in both Black and White rates.

conviction as an outcome, I find prosecution has a criminogenic effect (i.e., marginally prosecuted individuals are more likely to be re-arrested or re-convicted) that is more pronounced for Black individuals than White individuals. This finding suggests reducing prosecution (e.g., dismissing or diverting more) would lower future criminal involvement for both Black and White individuals, but more so for the former. This observation has salient policy implications that echo the results found in recent literature (Mueller-Smith and Schnepel, 2021; Agan et al., 2021).

However, although important to policymakers, these results miss a key detail: future convictions *require* future arrests. To this end, I develop a simple—yet novel—model to identify the causal effect of an additional conviction on the unconditional probability of future convictions (i.e., accounting for patterns in future re-arrest). I find that whereas an additional conviction increases the probability of a future conviction more for Black individuals, this effect is driven by future arrest patterns. Thus, after taking into account patterns in future arrests, additional convictions increase the unconditional probability of conviction more for White individuals. This result indicates prosecutors view signals—such as prior criminal convictions—through a racial lens and reiterate previously discussed results with respect to how prior convictions map into incarceration lengths.

All in all, my findings portray prosecutors as complex actors whose actions are informed by upstream forces in the criminal justice pipeline. For example, I find prosecutors use prior convictions to increase both the probability of conviction and carceral length, but less so for Black individuals relative to their White peers. These forces contribute to the observation that Black individuals are more likely to have their cases dismissed and receive shorter carceral sentences than similarly situated White individuals. Yet, these channels do not equalize outcomes for marginal defendants across racial lines. One potential mechanism that could explain these results is patterns in the selection process that bring people into the prosecutor's office, as well as institutional norms that constrain prosecutorial decisions

7

(Small and Pager, 2020). Specifically, if police arrest Black individuals at higher rates than White individuals—and one assumes the marginal-case strength/conduct decreases with the number of arrests per capita—this finding would drive a wedge between the marginal-case strength/conduct of Black and White arrestees. In this example, it would be plausible that a higher percentage of Black cases would be dismissed and the marginal case against a Black individual to be of lower strength/for lower conduct.

At the end of this paper, I attempt to provide some empirical support to validate this mechanism. First, I show massive racial disparities in the per-capita arrest rates in the jurisdiction of interest, whereby arrests per capita for Black men are roughly five times that of White men. Theoretically, this disproportionality can be driven by either Black individuals committing criminal acts at much higher rates (or in a manner that increases the probability of detection) and/or policing strategies that increase apprehension probabilities for Black individuals compared with White individuals who are engaged in similar behavior. To empirically test whether disparities in the level and nature of conduct can rationalize the disparity in arrests, I focus on drug crimes, because they uniquely enable me to proxy behavior.[14] I utilize a combination of machine learning, survey data on drug behavior, and proprietary cell phone data to construct proxies for conduct relating to illicit drugs and foot traffic at the census-block group level. I use these data to test if racial composition of an area correlates with the per-capita number of arrests made of residents in that area. I find a strong positive correlation with the percentage of the population of a block group that is Black and the per-capita drug arrests of residents living in said block group. This result suggests the selection process that brings individuals into the prosecutor's office may be biased in a way that drives the marginal Black arrestee to be drawn into the system with either a weaker case and/or for lesser conduct than White individuals. This highlights a potential source of upstream bias that likely influences prosecutorial decision-making.

---

[14]Cases involving drug crimes also exhibit the largest raw disparities in incarceration and represent the largest category of charges in my sample.

2 Related Literature

It is an exciting time to study the criminal legal system, and more specifically, prosecutorial decision-making, because it is an active literature with high-quality researchers in a multitude of different disciplines. The following section touches upon the existing literature that the current work aims to contribute to through the eyes of an economist. As such, this summary is far from exhaustive. I start with the seminal work of Rehavi and Starr (2014), who documents raw racial disparities in the length of federal carceral sentences. Similar to my findings, the authors note most of the raw disparity can be attributed to racial differences in criminal history and arrest charges. More recent work by Tuttle (2021) also examines prosecutorial behavior at the federal level. He uses a difference-in-bunching approach and changes in federal sentencing guidelines to study prosecutorial discretion regarding drug-charge decisions—again focusing on the use of mandatory minimums. The results in Tuttle (2021) suggest racial bias—in the form of shifting drug weights to trigger mandatory minimums—disproportionately increases sentence lengths for Black and Hispanic defendants relative to their white contemporaries. Therefore, in both Rehavi and Starr (2014) and Tuttle (2021), the results suggest changing statutory guidelines—in the form of mandatory minimums—would greatly reduce racial disparities in incarceration lengths.

This article uses data from a local prosecutor's office, whereas both Rehavi and Starr (2014) and Tuttle (2021) focus on federal prosecutors. Recent studies also leverage data from local prosecutors to explore prosecutorial behavior (Robertson et al., 2019; Agan et al., 2021; Harrington and Shaffer, 2021; Agan et al., 2022; Sloan, 2022; Mitchell et al., 2022; Shaffer, 2022). Shaffer (2022) uses statewide court data to show local prosecutors in North Carolina (NC) reduce racial disparities by penalizing Black defendants less than White defendants for prior criminal convictions. The author then links these data to a rich self-administered survey of prosecutors and shows the discount on prior convictions for Black relative to White defendants is likely driven by prosecutors' awareness of racial bias in the system.

These results are similar to those of the current paper documenting how local prosecutors use their discretion to mitigate raw racial disparities driven by upstream systemic biases.[15] Similarly, Harrington and Shaffer (2021) uses discontinuities in NC's sentencing grid, which determines mandatory prison time, to study how prosecutors use discretion in the form of charge reductions (i.e., reducing the severity of an arrest charge). The authors find convincing evidence that prosecutorial discretion has evolved over time to reduce racial disparities in arrests and that this evolution is concentrated in charges that were more likely to be initiated by police (instead of citizens). Again, these results are in agreement with the findings in the current paper that suggest *modern* prosecutors reduce racial disparities that are driven by upstream bias.

Mitchell et al. (2022) uses across jurisdiction variation to analyze the correlation between progressive chief prosecutors and racial disparities. They find jurisdictions that elect progressive chief prosecutors are less likely to exhibit Black-White racial disparities in case outcomes. Although causally attributing these patterns to prosecutorial discretion is difficult, they are consistent with results in this paper that show prosecutors potentially use their discretion to correct for upstream biases.

The extensive-margin analysis in my paper is most similar to Agan et al. (2021); Sloan (2022), who exploit quasi-random assignment of cases to prosecutors. The former is set in Suffolk County, Massachusetts, a jurisdiction roughly twice the size of ours, with a less conservative voting base, whereas the latter is set in New York City, which is also significantly larger and more liberal than the jurisdiction of interest. Agan et al. (2021) find the prosecution of low-level misdemeanors has a criminogenic effect. Specifically, the authors find that declining to prosecute these misdemeanors results in a 33-percentage-point decrease in the likelihood of a subsequent criminal complaint within two years (90% relative to the non-prosecuted mean). This finding is strikingly similar in magnitude to the

---

[15] Also note the setting in Shaffer (2022) is similar to that of the current study (i.e., a southern jurisdiction) and that they find prosecutors using discretion to reduce racial disparities may be a recent phenomenon, and my results come from recent (2015-2019) data.

31-percentage-point (100% relative to prosecuted mean) decrease in the likelihood of re-arrest caused by dismissals I find in this paper for non-violent "low-level" felonies. Sloan (2022) uses quasi-random assignment of cases to prosecutors to identify the effect of being assigned a prosecutor of a different race on conviction rates. The author finds being assigned a prosecutor of a different race increases the likelihood of conviction but only for property crimes. However, another interesting result from this paper is that being assigned a prosecutor of a different race actually *decreases* the probability of receiving a carceral sentence and *lowers* the average sentence length.[16] Both results are congruous with the results of this paper, whereby Black defendants receive significantly shorter sentences than similarly situated White defendants.[17]

My results on the inferred propensity to use/sell drugs are motivated by the existing literature focused on modeling what demographic characteristics drive police surveillance. Lum and Isaac (2016) and Bates et al. (2019) most closely inform the methodology implemented in this paper. Both utilize synthetic populations (e.g., publicly available census microdata assigned to small geographic areas) and models of drug use based on the National Survey on Drug Use and Health. The latter uses a logistic model to predict the prevalence of drug use in certain geographic areas to inform policies aimed at curtailing the opioid epidemic. Similar to the current paper, Lum and Isaac (2016) finds that in Oakland, California, drug use is uniformly distributed by race; however, drug arrests are concentrated in places where minority groups reside. Specifically, I find the racial composition of an area is strongly correlated with the number of drug arrests of people who reside in that area, even after accounting for population and the intensity of drug use/sales. This result echoes the existing literature that also finds Black individuals in the U.S. are more likely than their White counterparts to be arrested for drug crimes. For example, Mitchell and Caudy (2015) uses the 1997 National Longitudinal Survey of Youth (NLSY97)—a multiwave panel study

---

[16]Neither result is precisely estimated and statistically insignificant at the conventional 5% level with p-values of ~ 0.12 and ~ 0.63; thus, one should interpret the results with caution.

[17]Note the vast majority ($>90\%$) of prosecutors in my sample are White.

of a nationally representative sample of individuals between the ages of 12 and 16 as of December 31, 1996, in the U.S.—to study if racial disparities in drug arrests can be explained by "race differences in drug-offending, non-drug-offending, and/or neighborhood contextual features". The authors find only 15% of the racial disparities in drug arrests can be explained by differences in drug-use, drug-sales, non-drug-offending, or neighborhood features. They conclude these results—such as those contained in this article—are consistent with outcomes produced by racially targeted policing practices. Similarly, Beckett et al. (2005, 2006) find an over-representation of Black and Hispanic individuals in drug arrests in Seattle, Washington, for drug possession and sales, respectively. Both articles conclude that after controlling for racial differences in the intensity and nature of drug use and sales, unexplained racial disparities that are likely driven by policing strategies remain.[18]

The results contained in this paper provide a cautionary tale for policymakers looking to implement "color-blind" policies. These policies are motivated by the intuition that if decision-makers are unaware of race, it will, by definition, not factor into their decisions (Sah et al., 2015). As shown by a recent ethnographic study in Florida, "color-blinding" is also the expressed preference of most line attorneys (Dunlea, 2021). However, removing race as an input assumes all other inputs in the decision-making process are themselves "color-blind". An assumption that is unlikely to be true as contemporaneous and historical biases generate racial disparities in characteristics that have considerable influence over prosecutorial decisions (Doleac, 2022). For example, in interviews conducted by Dunlea (2021), prosecutors explicitly mention how Black defendants likely have lengthier criminal records, due to racially biased historical criminal justice interventions that act to disproportionately increase punishment for Black individuals today.

Until recently, this notion was only theoretical for prosecutors, due to technological limitations. However, the Stanford Computational Policy Lab (SCPL) recently developed

---

[18]Also, a large literature on biased policing outside the context of drugs exists. See, for example, (Anwar and Fang, 2006; Knowles et al., 2009; Rios, 2011; Goncalves and Mello, 2017; West, 2018).

a robust redaction algorithm "which automatically identifies and redacts race-related words from incident narratives" (Chohlas-Wood et al., 2021). Unfortunately, likely due to sample-size restrictions, a preliminary pilot was unable to detect statistically significant effects of redaction on charging decisions. Furthermore, no test of how redaction affected the average length of sentences was conducted.

3 Institutional Background

In the jurisdiction, prosecutors are referred all arrests for crimes with a potential penalty of more than 30 days of incarceration or a fine of $500 or more. Charges are referred to the prosecutor's office by law enforcement after an arrest and, for most charges, after an initial bond hearing. Once charges are referred to the office, they are put into the office's Case Management System with all necessary documentation before being assigned to a prosecutor. Charges are prosecuted "vertically" by the office, meaning the initial prosecutor assigned to the case will review the case for legal sufficiency, prosecute the case if necessary, and ultimately dispose of the case. This process is in contrast to "horizontal" structures where different prosecutors perform different parts of the process. Defendants facing multiple charges stemming from the same incident or set of facts generally have those charges bundled together into a single case. Multiple charges on the same defendant from separate incidents may be prosecuted separately; regardless, the same prosecutor often handles all of the charges for a single defendant.

The prosecutor then presents the charge(s) at a preliminary hearing to determine whether the probable-cause threshold for an arrest has been met. At the preliminary hearing, the charge(s) may be (1) remanded (i.e., sent to) to municipal court, (2) dismissed by the judge or prosecutor, or (3) "bound over", meaning the probable-cause threshold has been met and the prosecutor proceeds with prosecution. If a charge is bound over, the prosecu-

tor presents the charge to the grand jury via an indictment. The grand jury then acts on the indictment to either "true bill" the indictment, allowing the prosecution to continue, or returning a "no bill", ending the prosecution. After an indictment is true billed, the charge will be disposed of via trial, plea, or dismissal, with the former being rare.

Importantly, I observe the charges at each pivotal stage:

- At Arrest/Bond Hearing - the initial charges as determined by law enforcement.

- At Indictment - the charges that the prosecution believes it can prove beyond a reasonable doubt. In certain rare instances, prosecutors issue a direct indictment such that the individual is indicted pre-arrest.

- At Disposition - the charges the defendant is ultimately found guilty of, not guilty of, or pleas to, or that are dismissed.

This level of granularity allows me to measure how prosecutors alter defendants' charges as they progress through the prosecutorial process. This ability is especially important for understanding the interaction between charges for a given defendant, because each charge is not prosecuted in a vacuum. Rather, all actors in the criminal legal system—prosecutors, defendants, judges, and defense attorneys—are considering all charges a defendant faces during prosecution. Concurrent charges affect what charges are initially filed, plea negotiations, and likely even sentencing. Given this reality, the analysis can be done at the charge, case, or defendant-case level. I create the defendant-case level by grouping all charges for the same defendant with the same arrest date or that were disposed within five days of each other.[19] In general, charges with the same arrest date or with disposition dates near in time were probably prosecuted as a whole. Creating this defendant-case-level index allows me to properly relate all charges that are dismissed as part of larger plea negotiations—one of the

---

[19]The total number of defendant cases remains almost constant, varying as I vary this window from 3 to 14 days.

most common ways charges are disposed. For brevity, I refer to defendant cases as simply cases henceforth.

Importantly, sentencing data are at the charge level for all charges resulting in a guilty verdict—plea or trial. The sentencing data present a number of interpretation challenges, but I have chosen to break down sentences into four components:

- Incarceration - the number of days served or to be served in jail or prison

- Exposure - the maximum number of days served or to be served in jail or prison

- Probation

- Fine

Throughout my analysis, I focus on the incarceration of the top charge at the case level. The vast majority of sentences in the jurisdiction are concurrent at the case level; thus, most defendants only face the incarceration punishment of the top charge.

Most cases that are prosecuted result in a guilty plea. Because the prosecutor negotiates the plea agreement with the defendant, including determining which charges the defendant is guilty of, the prosecutor plays a significant role in the defendant's punishment. Since punishments for a given crime are dictated by statute, the crime to which a defendant pleads dictates the range of possible punishments. Once the plea is settled, it is presented to the judge and a punishment is imposed. Determining the punishment can be challenging, because prosecutors may make recommendations about the punishment, including recommendations the defendant has agreed to. Although a judge often has the authority to set the final punishment, the prosecutor significantly influences this decision through any recommendation, and most judges simply impose the recommendation. To offer empirical support for this assertion, I examine how much of the variation in actual sentences can be explained

by recommendations made by prosecutors.[20] I find recommendations explain the majority of variation in actual incarceration ($R^2 = 0.66$) and are highly correlated ($\beta^{OLS} = 1.1$, robust s.e.=0.04).[21]

4 Data and Descriptive Statistics

4.1 Administrative Data

The jurisdiction provides weekly data dumps consisting of multiple tables from the jurisdiction's Case Management System (CMS) database. The jurisdiction's database manager queries the data based on an agreed-upon query and uploads the data to a secure file-storage container. The data are then passed through an extensive data-cleaning process that combines tables, creates new columns, and outputs a large flat file, where each row is a single charge. Part of the cleaning process includes identifying data quality issues and presenting them to the jurisdiction to be addressed. As such, the data are validated by the jurisdiction for accuracy.

This section will present broad descriptive statistics at both the charge and the case level using the data generated from the processes described in section 3. To begin, looking at the entire sample as a whole, before any filtering takes place, is instructive.[22] Table 5 presents the breakdown of all charges from March 1, 2015, to current by race, gender, age, outcome, time of arrest, number of victims, number of co-defendants, and type of defense counsel. I apply some filtering to narrow down the sample. First, I remove any charge that was disposed of after the normal operation of the court system was interrupted due

---

[20]I only observe prosecutorial recommendations regarding incarceration for cases disposed after January 2019, because the office began collecting data that month on recommendations made.

[21]Using quantile regression methods, I also observe that recommended sentences lengths correlate highly with actual sentences throughout the distribution of incarceration, with the coefficient ranging from 1 to 1.3 from the $20^{th} - 80^{th}$ quantiles, respectively.

[22]I dropped 285 charges that were identified by the office as errant data entry. These included missing gender, missing sentences, and duplicate warrant numbers (charges).

to COVID-19.[23] I also limit to Black and White males aged 18-65. The latter filter is to avoid non-working-age elderly as well as youth offenders, who may be treated differently. The former is to isolate what I believe to be the two largest comparison groups where I may expect the largest amount of racial bias. Finally, I remove charges that have an outcome of "Other"; the majority of these charges were either remanded to a lower level municipal court or failure to appear charges. These filters leave me with 42,761 charges. The final filter isolates the subset of charges that are quasi-randomly assigned. To do this, I drop any charge that is part of a case that has any charge that is categorized as involving domestic violence and/or violent by state statute. These charges include, for example, cases involving murder, attempted murder, manslaughter, kidnapping, and armed robbery with a deadly weapon.[24] After applying these filters, I am left with 18,972 charges. Table 1 illustrates the type of charges that make up my sample, with the total charges and percentage of total for broad categories as well as the top three charges within the category. The majority are charges involving low amounts of drugs, traffic, unlawful possession of guns, and low-value property.[25]

After collapsing to the case level, I am left with 11,704 cases involving 9,530 individuals. Table 2 is a balance table showing the stark differences in cases by race. The first row of the table illustrates the raw disparity of 37 days in average incarceration. Another notable difference that I highlight later in the paper is in the criminal history and the number of charges per case. Specifically, Black men have no prior convictions in 32% of cases in my sample; conversely, nearly half of the White men (49%) have no prior convictions. For the latter, 66% of White defendants face single-charge cases, whereas only 59% of Black men do. Black men are slightly more likely than White men to be charged with crimes that do not

---

[23]I conservatively set the date of this to March 1, 2020.

[24]My results regarding the decomposition of disparities in incarceration are robust to the inclusion of these charges; however, these charges are not quasi-randomly assigned.

[25]Some charges, however, are "person" crimes including, for example, assault and battery 1st/2nd degree; the state does not statutorily consider these crimes violent offenses, and such offenses do not include serious or grave injury. These offenses include mutual fights, shoving, and threats of violence.

Table 1: Composition of charges in the sample of interest

Category	Subcategory	N	%	Category	Subcategory	N	%
Drugs		**7,315**	**39**	**Traffic**		**2,117**	**11**
	Poss <1 g Meth/Crack, 1st Offense	*1,267*	*17*		*Failure to stop for blue light*	*818*	*39*
	PWID Non-marijuana, 1st Offense	*1,234*	*17*		*Driving under suspended license*	*509*	*24*
	MDP Non-marijuana, 1st Offense	*847*	*12*		*DUI, <.10, 2nd Offense*	*232*	*11*
Property		**4,956**	**26**	**Person**		**1,103**	**6**
	Shoplifting <$2000	*755*	*15*		*Assault and Battery 2nd Degree*	*430*	*39*
	Breaking into Motor Vehicle	*591*	*12*		*Assault and Battery 1st Degree*	*150*	*14*
	Forgery <$10,000	*393*	*8*		*Indecent exposure*	*92*	*8*
Guns		**2,513**	**13**	**Other**		**968**	**5**
	Unlawful carrying of pistol	*1,447*	*58*		*Resisting Arrest*	*238*	*25*
	Unlawful sale/delivery of handgun; stolen handguns	*588*	*23*		*Assault/Beat/Wound Police*	*156*	*16*
	Possesion of firearm/ammo by person convicted of violent felony	*237*	*9*		*Threatening Life/Person/Family of Public Official/Teacher/Principal*	*108*	*11*

Note: PWID: Possession with intent to distribute; MDP: Manufacturing/Distribution/PWID; DUI: Driving under the influence

involve a victim (59% vs. 56%). They are also more likely to resolve cases through a plea bargain, and for at least one charge, to have a carceral sentence, though they receive one week less of time served on average. Finally, cases involving Black men take 39 days longer to dispose, regardless of the outcome of the case.

My data allow me to pinpoint the exact locations of both the arrest and residence of each arrestee in my sample. The raw data provided list two lines of street address, city, state, and zip code for each arrest, as well as the residence of the arrestee. I take this information and utilize the python package *geocoder*[26] to determine the latitude and longitude of each address. This information is then mapped over Census shapefiles to place each arrest and residence inside a census block. This approach allows me to merge into my administrative data a multitude of census features, as well as create proxies for the propensity to use and/or sell drugs.

---

[26]Documentation found here.

Table 2: Balance table illustrating differences between cases involving Black and White male defendants across multiple categories

	(1)	(2)	(3)
Variable	White	Black	Difference
Incarceration (days)	253.0	290.4	37.4**
Probation (days)	334.7	336.0	1.2
Fine amount ($)	52.9	36.3	-16.6**
Incarceration rate (%)	47.8	50.0	2.2**
Case dismissal (%)	25.8	26.1	0.4
Plea rate (%)	60.1	64.6	4.6***
Time to disposition (days)	324.1	363.2	39.1***
Time served (days)	67.5	58.7	-8.8***
Criminal History:			
0	49.2	32.3	-16.9***
1	13.5	13.5	0.1
2	9.2	10.7	1.4**
3	6.4	9.1	2.7***
4	4.7	7.0	2.4***
5	3.8	5.3	1.5***
6	2.9	4.6	1.7***
7	2.4	3.1	0.7**
8	7.9	14.3	6.4***
Victim count:			
Victimless	55.5	59.4	3.9***
Single victim	34.5	32.3	-2.2**
Multi-victim	10.0	8.3	-1.7***
Number of Charges:			
1	65.7	59.1	-6.7***
2	19.0	21.7	2.7***
3	7.0	9.3	2.3***
4	8.2	10.0	1.8***
Observations	4,943	6,761	11,704

Note: For simplicity, criminal history and the number of counts were truncated such that the last bins encompass all values equal to and above the 90th percentile value.

4.2 Drug Use/Sales

To estimate the prevalence of drug use and sales within the jurisdiction, I turn to the National Drug Use and Health Survey (NSDUH). This national survey has been run annually since 1971 by the Substance Abuse and Mental Health Services Administration (SAMHSA), an agency in the U.S. Department of Health and Human Services (HHS). It uses a unique face-to-face and computer-assisted hybrid interview approach to obtain a true response to

sensitive questions about substance abuse and mental health. The survey samples $\sim 70,000$ respondents from all civilian non-institutionalized populations, ages 12 or older, in the 50 states plus DC. A simple Google Scholar search for "National Drug Use and Health Survey" returns more than 4 million results. Policymakers have also cited the NSDUH when crafting responses to the recent opioid crisis (McCance-Katz et al., 2017), as well as in reports released by the Surgeon General (Office of the Surgeon General U.S. Department of Health and Human Services , HHS).

However, as with any survey that claims to represent the general population, some caveats are in order. A recent criticism of NSDUH by Reuter et al. (2021) presents four "sources" errors that can cause NSDUH to severely underestimate drug use and sales. I go through each critique and discuss how it can impact my analyses that aim to estimate the racial gap in drug use and sales in the Appendix B. To summarize, I believe my analysis— although not immune from all the critiques in Reuter et al. (2021)—suffers less than most applications of the NSDUH, because my aim is to measure differences between racial groups and not levels of drug behavior.

I focus on questions related to recent non-marijuana and marijuana drug use and general drug sales.[27] As Table 3 shows, drug use and sales are low-probability events. Another salient pattern from Table 3 is that Black males are less likely to use all drugs except marijuana and crack cocaine and are more likely to sell drugs. The survey also contains a rich set of demographic questions that I use to predict the probability of drug use within high-quality census microdata sampled from the Public Use Microdata Areas (PUMA) covering the same geographic area as the jurisdiction. The dimensions I use are age, marital status, gender, education, race, employment status, personal income, family income, and household income relative to the federal poverty line.

---

[27]I divide drug use into non-marijuana and marijuana, because the jurisdiction treats marijuana much more leniently.

Table 3: Weighted response rates (%) for recent activity (≤1 year) for males by race

	Non-hispanic Black	Non-hispanic White	Ratio
All non-marijuana	7.0	9.8	0.7
Any sales	2.7	1.9	1.4
Marijuana	19.6	16.9	1.2
Crack	0.82	0.35	2.3
Cocaine	1.8	2.0	0.9
Heroin	0.29	0.46	0.6
Hallucinogens	1.4	2.2	0.6
Inhalants	0.50	0.58	0.9
Methamphetamine	0.43	1.2	0.4
Prescription pain-killers*	3.6	5.0	0.7
Prescription tranquilizers*	1.5	2.7	0.6
Prescription stimulants*	0.79	2.4	0.3
Prescription sedatives*	0.23	0.50	0.5

Note: Data come from the NSDUH subset to male respondents from core-based statistical areas with fewer than 1 million people classified as "metropolitan" (based on USDA 2013 Rural/Urban Continuum Codes). "Any non-marijuana" includes all rows except "Any sales" and "Marijuana".
*Only includes misuse (i.e. not taken as prescribed)

5 Racial Bias in Incarceration

This section provides descriptive evidence of how prosecutors influence racial disparities in incarceration lengths. I first explain the empirical methods used to analyze the factors that contribute to total disparities in the length of incarceration. To organize my findings, I utilize the theoretical framework constructed in Bohren et al. (2022). They show—using a decomposition inspired by Kitagawa (1955), Oaxaca (1973), and Blinder (1973)—that total discrimination is a combination of direct and systemic components. In my setting, total disparities represent an important and policy-relevant benchmark. After accounting for a rich set of legally permissible features, we define any residual correlation with race to be driven by direct bias by prosecutors. As defined in Bohren et al. (2022), the difference between the total and direct components represents the systemic component. In our context where all of the observables we condition on are pertinent to the legal system (e.g., prior criminal convictions, the exact arrest charges, the number of charges in a case, etc.), this

component can be understood as a "legally" systemic component. I then describe and apply a novel decomposition method developed by Gelbach (2016) that enables me to identify how much each characteristic contributes to the total systemic bias.

5.1 Framework

5.1.1 Total Disparity

First, I set out to benchmark the total racial disparity in incarceration, by estimating the following regression model:

$$y_i = \alpha + \beta_1 R_i + \theta X_i + \epsilon_i \ , \tag{1}$$

where y_i is the length of the carceral sentence for the defendant i, R is an indicator for whether the defendant i's race is recorded as Black, X_i can be a vector of observables that the econometrician chooses to remove from the subsequent decomposition, α is a constant, and ϵ is the idiosyncratic Gaussian error term. Unlike in most settings, the raw disparity (i.e., β_1, where $X_i = \emptyset$) is policy relevant, because it is a simple measure of the amount of total punishment imposed by race.[28] However, to remove all demographic information from the subsequent analyses, I include the age of the defendant at arrest in X_i (note I implicitly control for gender because I only compare across the male population).[29]

5.1.2 Direct Disparity

My main objective is to understand how prosecutors use their discretion to influence outcomes. This section details how I attempt to get an estimate of direct bias, which I assume

---

[28]In fact, one of the first questions the lead prosecutor asked was whether a raw racial disparity existed in the length of carceral sentences.

[29]Conditioning on age does not substantively change my results (see raw measure in Table 2); it is merely my preference to remove all available demographic information from the potential systemic channels.

is mainly influenced by prosecutorial discretion. Ideally, an experiment would be conducted in which defendants were randomly assigned race after arrest and then incarceration values were observed. Therefore, any difference in incarceration could be causally attributed to race. For obvious reasons, this experiment is impossible to run.[30] Instead, I leverage high-quality administrative data that allow me to compare Black and White men who are the same age at arrest, with the same number of prior criminal convictions, the same arrest-charge code (e.g., identifying the crime being alleged by law enforcement), the same number of charges within their case, the same number of victims, the same number of co-defendants, and the same type of defense counsel (e.g., public vs. private vs. no counsel with N number of cases seen in my sample). I then assume any residual variation explained by race is attributable to prosecutors treating similarly situated (in a legal sense) defendants differently due to race.[31] This assumption fails if any unobserved legally permissible variables are correlated with cases involving Black defendants and the residual variation left over after conditioning on the aforementioned set of controls. Operationally, I estimate the following model using OLS:

$$y_i = \alpha + \beta_2 R_i + \boldsymbol{\gamma C_i} + +\theta X_i + \epsilon_i \ , \tag{2}$$

where all variables are the same as in equation 1 with the addition of a vector of controls, $\boldsymbol{C_i}$ described in the previous paragraph. Therefore, β_2 is an estimate of the direct effect of race on incarceration, which is attributable to prosecutors treating similarly situated defendants differently due to race.

5.1.3 Systemic Disparity

The difference in total disparity and direct disparity represents the systemic—or upstream—disparity in the system. Exploring what features contribute to systemic disparity is impor-

---

[30]See Greiner and Rubin (2011); Rose and Shem-Tov (2021), among others, for discussions on whether this experiment should be theoretically considered or would be possible.

[31]See section 3 for evidence of prosecutorial influence over carceral lengths.

tant, because this information can help one understand how policies that place weight on certain observables may affect racial disparities. For example, I find prior convictions drive the majority of total racial disparities in incarceration; therefore, any policy that places more weight on prior convictions will likely exacerbate the racial gap in incarceration. To investigate what drives systemic disparity, I turn to the decomposition method developed in Gelbach (2016). This method—similar to a Kitagawa (1955); Oaxaca (1973); Blinder (1973) decomposition—provides a clear strategy to decompose the sets of controls that explain the change from $\overline{\beta}_1 \rightarrow \beta_2$. To begin, I divide the vector of controls into K mutually exclusive categories:

1. Fixed effects for arrest charge

2. Fixed effects for number of prior convictions [32]

3. Fixed effects for number of charges within a case[32]

4. Fixed effects for max number of victims for a charge within a case[32]

5. Fixed effects for number of co-defendants in the case[32]

6. Controls for the type of defense counsel (e.g. indicators for private vs. public vs. none) as well as the number of cases handled by the defense counsel in my sample.

As one can see, each category can have multiple variables (e.g., the first category contains hundreds of arrest-charge fixed effects). For simplicity, I describe how to find the effect of each individual variable c and show that in the end, one can sum the effects of each c by group. First, I estimate the complete model with all included controls and recover the estimates $\hat{\gamma}^c$ for each variable c in the vector of controls X. Then, for each variable c, I run the following regression:

$$X_c = \Gamma^c \mathbf{X}_{-c} + \epsilon_c \,, \tag{3}$$

---

[32]To avoid running into bins with small samples when controlling for indicators, I hard code all values for each control that falls above the 90th percentile to the 90th percentile value.

where $-c$ represents all the variables in X without including c and save all coefficients, $\hat{\Gamma}^c$. Finally, I multiply $\hat{\Gamma}^c$ by $\hat{\gamma}^c$. This product represents the component of the change in $\overline{\beta}_1 \rightarrow \beta_2$ due to the inclusion of X_c as a control. As mentioned above, these effects can be summed to obtain the total impact by subgroups, K.

$$\text{Impact of controls in subgroup } K = \sum_{c \in K} \hat{\Gamma}^c \cdot \hat{\gamma}^c \,. \tag{4}$$

Note I interpret β_1 as an estimate of total bias, β_2 as an estimate of direct bias introduced by the prosecutorial discretion, and $\beta_1 - \beta_2$ as an estimate of systemic bias. Therefore, the decomposition allows me to recover the effect of each group of controls on systemic bias.

5.2   Results

Figures 1a and 1b provide point estimates of total and direct disparities, using all cases and only cases with at least one guilty charge, respectively. In both cases, I identify large and statistically significant total disparities that represent approximately 15% of the mean incarceration. Specifically, I find that, on average, compared with White men of similar ages, Black men are sentenced to carceral sentences that are 15% longer. I then include indicators for the number of prior criminal convictions, the arrest-charge code of the top charge (e.g., identifying the crime being alleged by law enforcement), the number of charges within a case, the number of victims, number of co-defendants, the defense counsel type (e.g., public vs. private vs. none).[33] Finally, I include as a continuous variable the number of cases seen by the defense counsel in the sample as a proxy for defense counsel experience. Therefore, my estimate of direct bias is driven by any residual variation that can be explained by an indicator for whether the office identifies the defendant as Black. Again, I find similar results in both the full sample (Figure 1a) and the sample excluding case dismissals (Figure 1b) in

---

[33]I add each control as an indicator to allow for non-linearity in the effect of each variable on average incarceration.

which Black men receive sentences that are 13% shorter than similarly situated (with respect to all the observables listed above) White men. The symmetry of my estimates around zero implies large systemic (or upstream) features exist that work to disproportionately increase sentence lengths for Black men relative to White men. These results also suggest prosecutors use their discretion to decrease the average sentence for Black men.[34]

To preview my results in the next section, I find the observable that drives systemic bias is the number of prior convictions. Figure 2c illustrates the marginal effect (i.e., the effect holding all other controls in the model constant) of additional prior convictions on carceral lengths. For both Black men and White men, additional convictions increase sentence lengths; however, an interesting pattern becomes salient: the initial increase associated with one to three convictions is significantly higher for White men. This finding is consistent with a model in which prosecutors have some race-specific expectation about the quality of information contained in the signal of prior convictions and therefore utilize this information differently by race. For example, assume prosecutors expect that—due to upstream biases— convictions were made with some race-specific probability of error M_R and $M_{Black} > M_{White}$ (e.g., if $M_{Black} = 10\%$ and $M_{White} = 5\%$, the implication would be that they expect a Black (White) individual with two convictions to have received them both "by mistake" with a probability equal to 1% (0.2%); because the probability of all priors being driven by mistakes becomes increasingly small for both races, prosecutors begin to use convictions as a signal of true criminality for both equally. Figure 2c implies that after three convictions, both Black and White individuals experience statistically indistinguishable penalties for their prior convictions (i.e., no difference exists between average incarceration for Black and White defendants with four or more prior convictions).

To guard against these results being driven by Black and White individuals having a

---

[34]The assumption here is that variation in sentence lengths are mainly attributable to prosecutors and not judges. As stated earlier, I believe this assumption is reasonable, because prosecutors wield the power to (1) prosecute or dismiss, (2) decide to upgrade or downgrade arrest charges, (3) plea bargain with defendants, and (4) make recommendations regarding sentence lengths to judges, which judges routinely accept.
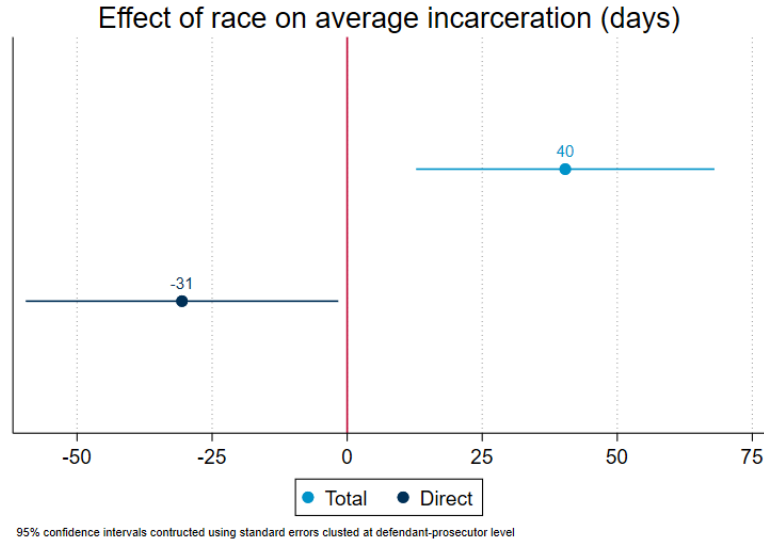
Effect of race on average incarceration (days)

(a) The sample includes all cases, cases that were entirely dismissed are included as 0 incarceration. The mean incarceration in this sample is 263 days.

Effect of race on average incarceration (days)
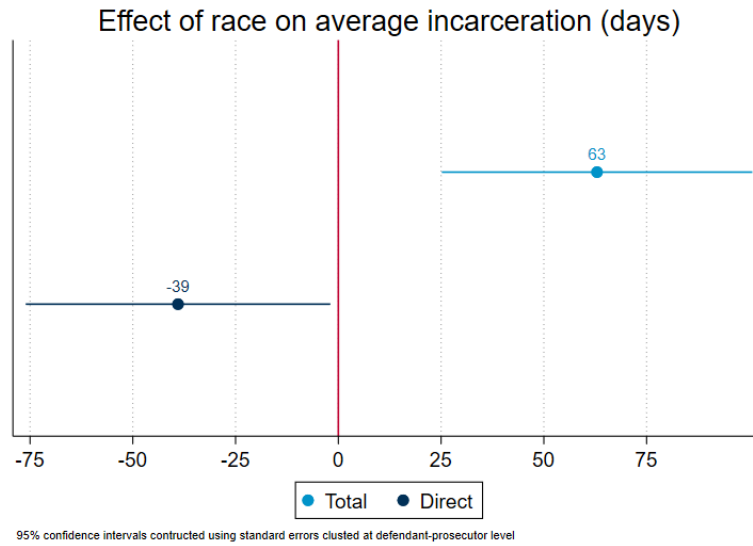
(b) The sample includes all cases; cases that were entirely dismissed are excluded. The mean incarceration in this sample is 359 days.

Figure 1: Total and Direct Bias in Incarceration

Note: Dependent variable in both (a) and (b) is the max incarceration of all charges within a case. Total estimates are conditional on the defendant's age at arrest, additional controls in direct models include number of prior criminal convictions, the exact arrest-charge code of the top charge (e.g., identifying the crime being alleged by law enforcement), the number of charges within a case, the number of victims, number of co-defendants, the defense counsel type (e.g., public vs. private vs. none), and the number of cases seen by the defense counsel in the sample.

different portfolio of prior convictions, I first show in Figure 3 that no large racial differences exist in the composition of priors for individuals with a certain number of prior convictions. In other words, a Black individual with three prior convictions is as likely to have two prior felonies and one prior misdemeanor as a White individual with three priors. Furthermore, I show the marginal effect on incarceration is higher for White than for Black individuals for one to two prior felonies (Figure 2a) and first non-felony (Figure 2b), with statistically insignificant differences thereafter. This result is strikingly similar to those documented in Shaffer (2022), who confirm by survey methods that prosecutors who believe total racial disparities in the system are likely driven in some part by bias (instead of disparity in conduct) use information from prior convictions less for Black individuals than for White individuals. This finding suggests that an explanation for my results is that prosecutors use race as a lens to interpret signals from criminal records.

In summary, I find a positive total disparity (i.e., Black defendants receive longer sentences on average) and negative direct bias (i.e., similarly situated Black defendants receive shorter sentences on average). These two facts imply a large systemic bias works to disproportionately increase sentences for Black men relative to White men. The estimates of total and direct bias in Figures 1a and 1b indicate systemic bias accounts for Black individuals receiving 28% longer sentences (with respect to the mean). To uncover what factors are contributing to this disparity, I utilize the Gelbach (2016) decomposition described in section 5.1.3.

Figures 4a and 4b illustrate the contribution of each set of controls to the total systemic disparity (i.e., the difference between the estimates of total and direct disparities). A positive (negative) number indicates the subset of controls decreased (increased) the gap between total and direct disparities, thus contributing (subtracting from) to the systemic disparity against Black individuals. The results suggest three major observables drive almost all systemic bias against Black defendants: the number of prior convictions, the number of

28
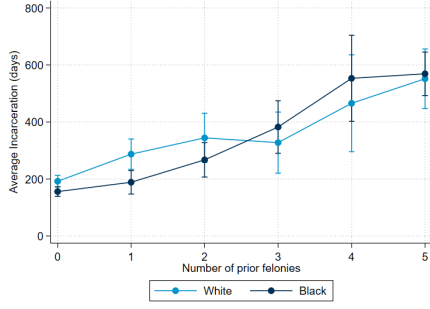
(a) Prior felony convictions
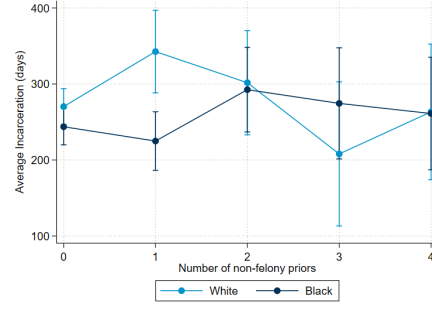
(b) Prior non-felony convictions
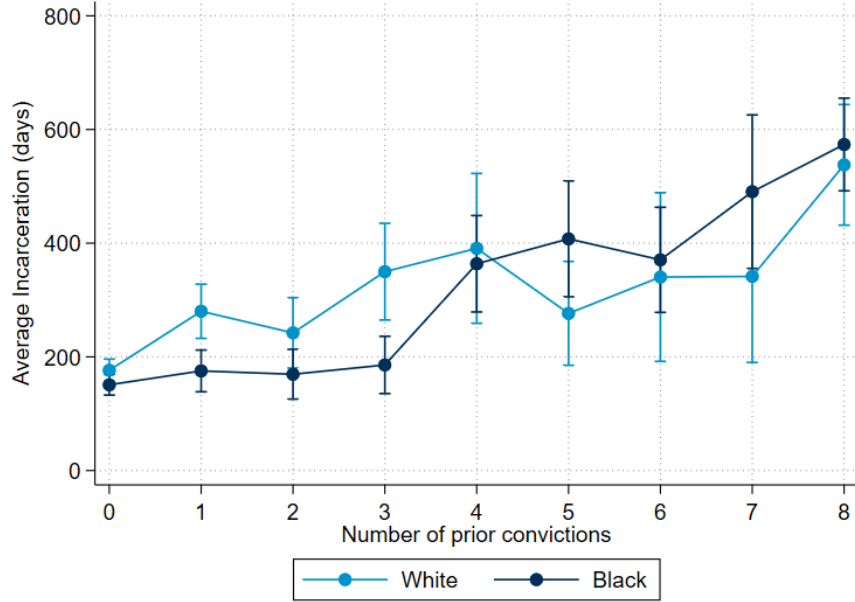
(c) All prior convictions

Figure 2: Marginal effect of additional prior (a) felony convictions (b) non-felony (either a misdemeanor or "unknown", 97% of the time this represents a misdemeanor) (c) total convictions on the average carceral length, for Black and White men.

charges per case, and the charge referred by the arresting officer. The first is significantly larger than the latter two. These results imply any policy—even if applied in a race-neutral way—that places weight on prior convictions will have large racially disparate impacts.

To further stress this point, I provide an example of how forcing prosecutors to use a race-neutral formula when charging could have large unintended consequences. For this example, I turn to the subset of my data in which individuals were accused of a crime
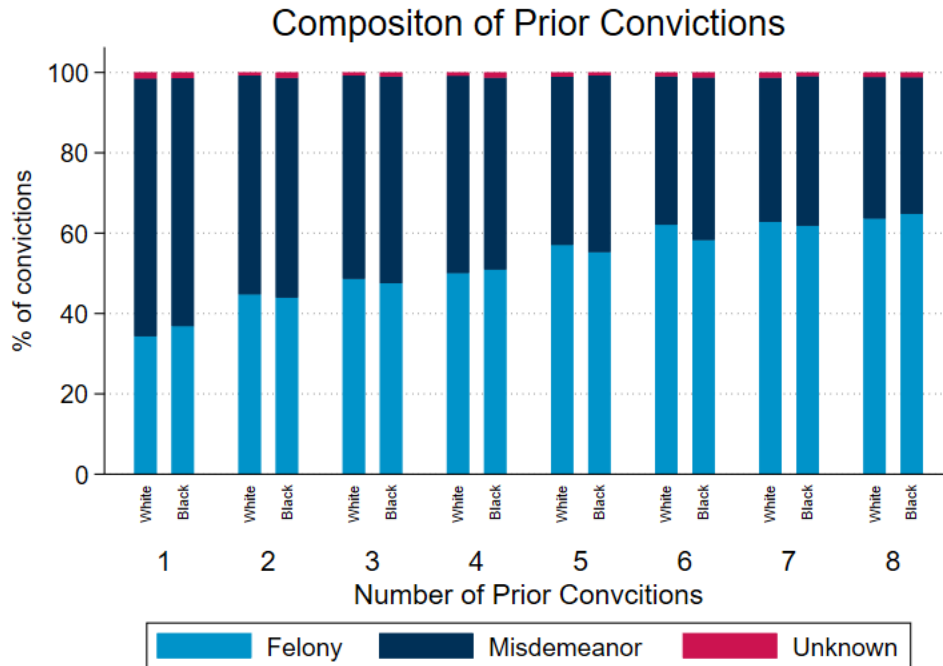
Figure 3: Composition of priors; each bar represents the percentage of cases with an individual with N priors that are felonies, misdemeanors, and unknown, by race
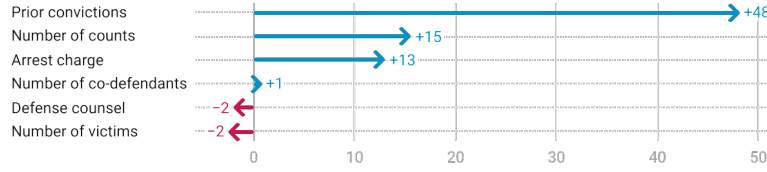
involving drugs. For all of these individuals, I was given access to PDFs of their rap sheets, which contain all criminal records available to prosecutors. I then scraped the rap sheets to create a database that allowed me to calculate a detailed criminal history of each individual.[35] Specifically, I was able to use each person's criminal history, the charge they were brought in on, and the state's statutory guidelines to determine what an individual was eligible to receive. I then coded a discount as anytime a person was convicted of a charge that was lower than their eligibility.[36] I find two interesting results: (1) Black individuals are much more likely to be eligible for enhanced drug charges (60%) than White individuals (30%); and (2) prosecutors discount a vast majority of the time, with around 80% of individuals receiving a discount, with no significant racial heterogeneity. Due to the former results, the latter discretion disproportionately helps Black individuals. If, instead, prosecutors were

---

[35]See Appendix F for a more detailed explanation of how I processed these rap sheets.

[36]For example, a discount would occur if someone came in with a possession-of-cocaine charge and had enough prior convictions to be eligible for a second offense, but was convicted of a first offense.
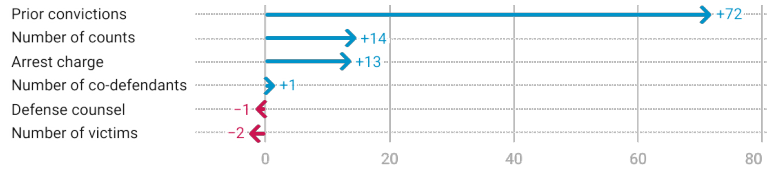
Contributions to Systemic Bias

Arrows represent how much (in days) each set of controls contributes to systemic bias, using
decomposition in Gelbach (2016)

Prior convictions	+48
Number of counts	+15
Arrest charge	+13
Number of co-defendants	+1
Defense counsel	−2
Number of victims	−2

(a) The sample includes all cases; cases that were entirely dismissed are included as 0 incarceration.
The mean incarceration in this sample is 263 days.

Contributions to Systemic Bias

Arrows represent how much (in days) each set of controls contributes to systemic bias, using
decomposition in Gelbach (2016)

Prior convictions	+72
Number of counts	+14
Arrest charge	+13
Number of co-defendants	+1
Defense counsel	−1
Number of victims	−2

For only cases that were not fully dismissed

(b) The sample includes all cases; cases that were entirely dismissed are excluded. The mean incarceration in this sample is 359 days.

Figure 4: Utilizing the Gelbach (2016) decomposition, both figures show the total number of days that each set of observables contribute to the total systemic bias.

unable to exercise discretion and were required to strictly follow charging guidelines, racial disparities would increase significantly.[37]

6 Racial Bias in Prosecution

At the extensive margin, prosecutors wield the largest amount of influence when deciding whether charges that are brought against individuals should be prosecuted.[38]. In our sample, if a charge is prosecuted, the defendant receives a conviction 99.81% of the time, with nonzero incarceration 70% of the time. To understand how prosecutors influence the relative

---

[37]The reason is that enhanced drug charges exponentially increase the length of sentences, and Black individuals are much more likely to be eligible for these enhanced charges.

[38]Prosecutors usually decide to not prosecute a case (around 87% of charges in my sample of interest), with the remaining cases being dismissed by judges.

distribution of people in the system and provide some context for the rest of my analysis, I estimate direct bias in dismissal decisions in much the same way I did in the previous section. I then leverage quasi-random assignment of charges to prosecutors to recover race-specific causal estimates of how non-prosecution affects future outcomes. Additionally, I utilize the theoretical model in section 6.1.2 to further decompose these causal estimates and identify how additional convictions affect the probability of conviction after accounting for arrest patterns.

6.1   Framework

6.1.1   Causal Effect of Non-prosecution

Identifying causal estimates presents an empirical challenge, because non-prosecution decisions are endogenous; thus, OLS estimates likely suffer from omitted-variable bias. For example, prosecutors may utilize unobserved features (e.g., tone in law-enforcement reports) to make prosecution decisions that are also correlated with the underlying "criminality" that affects future outcomes. In addition, OLS estimates are likely to suffer from infra-marginality bias (i.e., they fail to identify the effect on the margin).[39] To ameliorate these concerns, I utilize quasi-random assignment of charges to prosecutors and an instrumental-variable (IV) framework. The estimates represent the average effect for compliers.[40] Compliers, by definition, are marginal in the sense that they are the defendants on whom the prosecutors "disagree" on (i.e., some prosecutors would prosecute whereas others would not). Therefore, I interpret my results as identifying the causal effect of non-prosecution for defendants with the marginal charge. As noted in Agan et al. (2021), this LATE/CACE is also a policy-relevant treatment effect (PRTE) (Heckman and Vytlacil, 2001), because it identifies the effect of non-prosecution on the marginal charge by increasing the leniency of the prosecu-

---

[39]See (Simoiu et al., 2017; Arnold et al., 2018) for detailed explanations and examples.

[40]This effect is frequently referred to as the local average treatment effect (LATE) or complier average causal effect (CACE).

tors.

I first calculate the percentage of time that a charge is not prosecuted, either by dismissal or diversion (leniency henceforth), for each prosecutor.[41] I then residualize this leniency measure by prosecutorial team by time by broad crime-type fixed effects to isolate variation that is driven by quasi-random assignment of cases to prosecutors.[42] To avoid mechanically introducing correlation between charge characteristics and leniency, I calculate the leniency of a prosecutor for each charge using all the other charges the prosecutor handled that did not include the defendant in the current charge. Formally, I calculate the leniency of prosecutor, p, handling charge, c, for defendant, i as follows:

$$L_{cp} = \left(\frac{1}{n_p - n_p^i} \right) \left(\sum_{c \in n_p} D_{cp} - \sum_{c \in n_p^i} D_{cp} \right) , \tag{5}$$

where n_p, n_p^i, and D represent the total number of charges handled by the prosecutor p, the number of charges brought against the defendant i handled by the prosecutor p, and an indicator that is unity when a charge is not prosecuted, respectively. Relative leniency L measures for each prosecutor are then used as an instrument for non-prosecution in the following model:

$$F_i^t = \beta^t \hat{D}_i + \boldsymbol{\eta} \boldsymbol{X_i} + \tau + \epsilon_i , \tag{6}$$

where F is an indicator for a future outcome (e.g., re-arrest or re-conviction) for the defendant in the charge i t months after charge i is disposed, $\boldsymbol{X}$ is a vector of controls, and τ represents prosecutorial team by time by crime-type fixed effects. Formally, I use a two-stage least

---

[41]Operationally, I construct two leniency measures by prosecutor for each race.

[42]For all of the ensuing analysis, I use prosecutorial team by time by broad crime-type fixed effects, where time is defined as year-season-weekend and the broad crime categories include: Drugs, Property, Traffic, Guns, Person, and a catch-all "Other".

squares (2SLS) estimation strategy where the first stage can be written as

$$D_i = \zeta L_i + \boldsymbol{\gamma X_i} + \tau + \nu_i \ , \tag{7}$$

and is used to obtain causal estimates of $\hat{\beta}^t$ in the second stage (equation (6)).

To interpret my estimates as true LATE, I make four assumptions laid out in Imbens and Angrist (1994): random assignment, exclusion, relevance, and monotonicity. Here, I briefly discuss empirical and anecdotal support for each assumption, and Appendix C provides more details. Relevance is the easiest of the four assumptions to test empirically, for every IV regression, I include the Kleibergen-Paap F statistic (Kleibergen and Paap, 2006).[43] Conversely, random assignment is the hardest to empirically test. I rely on institutional knowledge that cases that do not include statutorily violent and/or domestic violence charges are assigned in a quasi-random way within prosecutorial teams.[44] Typically, these low-level cases come in and are assigned by managers using a first-in-first-out method to the "next person up".[45] Exclusion and monotonicity are also difficult to test directly; however, Frandsen et al. (2019) provides a joint test in which the null hypothesis is that exclusion and monotonicity hold, and I fail to reject this null over various parameter values of the test.

6.1.2 Decomposition Using Simple Model

I further decompose the LATE/CACE estimates by developing a simple novel model and using estimates of the average outcome for treated and untreated compliers using intuition from Abadie (2002) and empirical methods described in Angrist et al. (2013, 2022), which I describe in more detail in the Appendix E. The rest of this subsection describes the model

---

[43]One can test for relevance in multiple ways. I also utilize the *weakivtest* command in Stata, which implements the tests outlined in Olea and Pflueger (2013).

[44]I also guard against deviations from quasi-random assignment by further restricting to variation within teams by coarse crime type and within a certain time period (year-season-weekend).

[45]I also include an exhaustive set of controls that further guard against residual variation being driven by non-random selection. The fact that my results do not change substantively upon inclusion of these controls gives me additional confidence that I am utilizing quasi-random variation.

and its empirical predictions.

In this model, police screen (e.g., stop) individuals of race R_i for conduct $B_i(C_i)$ with probability $S(R_i, B_i)$ and, after screening, observe an individual's prior convictions C and decide whether to make an arrest with probability $A(B_i, R_i, C_i)$. I assume that the officers cannot observe C during screening. Therefore, the probability that an individual is screened and arrested, α, can be written as follows:

$$\alpha = Pr(\text{Arrest}|B_i, R_i, C_i) = S(R_i, B_i)A(B_i, R_i, C_i) \ . \tag{8}$$

After being arrested, prosecutors decide whether the facts of the case warrant a prosecution.[46] I assume prosecutors make this decision by setting some latent unobserved threshold above which they decide to prosecute. Without loss, assume some ordering U_i and re-scale the threshold to equal some constant that may depend on race γ^R. In this model, I assume this score is weakly increasing in the number of prior convictions, C, and conduct B. Therefore, the probability that an individual will be convicted can be written as follows:

$$\pi = Pr(\text{Conviction}|B_i, R_i, C_i, \alpha) = \alpha \underbrace{Pr(U_i \geq \gamma^{R_i})}_{\theta} \ . \tag{9}$$

Suppose the prosecution results in a conviction (i.e., very few charges that are prosecuted are found not guilty by trial[47]). I can then use equation 8 to define the causal effect of prosecution on the probability of future re-arrest, Δ_A as

$$\Delta_A = \underbrace{S(R_i, \tilde{B}_i)A(\tilde{B}_i, R_i, \tilde{C}_i)}_{\tilde{\alpha}} - \underbrace{S(R_i, B_i)A(B_i, R_i, C_i)}_{\alpha} \ , \tag{10}$$

where the first (second) term represents the probability of re-arrest for a marginally prosecuted (non-prosecuted) defendant, where the former has $\tilde{C}$ convictions and the latter has

---

[46]In this simple model, I lump all pertinent case facts into B.

[47]In the full sample of disposed non-dismissed charges, only 0.25% are found not guilty.

C convictions and $\tilde{C} > C$. Similarly, I can use equation 9 to define the causal effect of prosecution on the probability of future re-conviction, Δ_C, as

$$\Delta_C = \underbrace{\tilde{\alpha}\tilde{\theta}}_{\tilde{\pi}} - \underbrace{\alpha\theta}_{\pi} \ . \tag{11}$$

In this model, prosecutors have control over the probability of conviction, θ. Therefore, for the remaining analysis, I focus on θ by scaling the probability of conviction conditional on arrest, π, by the probability of arrest, α. I note that in this model, prosecutors have two channels to exercise discretion with respect to race:

1. **Race-specific thresholds**

$$Pr(U_i(C_i, B_i) \geq \gamma^R) \ \text{ such that } \ \gamma^B \neq \gamma^W \ .$$

2. **Race-specific effect of additional convictions**

$$\Delta\theta^B \neq \Delta\theta^W \ \text{ where } \ \Delta\theta^R = Pr(U_i(\tilde{C}_i, \tilde{B}_i) \geq \gamma^R) - Pr(U_i(C_i, B_i) \geq \gamma^R) \ .$$

Both channels can be written as functions of π and α, which can then be estimated separately for Black and White individuals using the following equations:

1. We can test for race-specific thresholds by testing whether:

$$\theta^B - \theta^W = \frac{\pi^B}{\alpha^B} - \frac{\pi^W}{\alpha^W} \neq 0 \ ,$$

or

$$\tilde{\theta}^B - \tilde{\theta}^W = \frac{\tilde{\pi}^B}{\tilde{\alpha}^B} - \frac{\tilde{\pi}^W}{\tilde{\alpha}^W} \neq 0 \ ,$$

2. While we can test for race-specific effect of additional convictions by testing whether:

$$\Delta\theta^B - \Delta\theta^W = \left[\frac{\tilde{\pi}^B}{\tilde{\alpha}^B} - \frac{\pi^B}{\alpha^B}\right] - \left[\frac{\tilde{\pi}^W}{\tilde{\alpha}^W} - \frac{\pi^W}{\alpha^W}\right] \neq 0 \ .$$

Writing out these tests, one notices the former requires comparisons *across* race of marginally prosecuted/non-prosecuted individuals. This comparison is problematic because it requires strong assumptions about how marginal charges vary *across* race. In other words, for a racial difference in θ to be truly driven entirely by a racial difference in γ, I must assume the individuals with marginal charges are identical across race, so that any difference in θ is driven by prosecutors setting more stringent thresholds. If, for example, marginal Black defendants have longer criminal histories, this fact would drive a positive difference between the race-specific θs. We know from Table 2 Black individuals have significantly longer criminal records than White individuals, so the marginal Black individual is likely to have a lengthier criminal history. For this reason, in section 6.2, I provide additional evidence that $\gamma^W > \gamma^B$ using a different source of variation, provided by a policy experiment.

The second test does not suffer from the aforementioned issues, because I first difference within race before comparing across race. I still need to make some assumptions to attribute racial disparity in $\Delta\theta$ to being the result of prosecutors using additional convictions differently by race. Namely, I must assume additional convictions do not affect conduct differently by race. If, for example, I find $\Delta\theta^B < \Delta\theta^W$, this observation could be driven by future conduct increasing more for White individuals than for Black individuals, $\tilde{B}^W - B^W > \tilde{B}^B - B^B$. If the converse is true ($\tilde{B}^W - B^W < \tilde{B}^B - B^B$), my results would still imply additional convictions are used to penalize Black individuals less than White individuals.

Using the empirical methods developed in Angrist et al. (2013, 2022) and described in Appendix E enables me to estimate all the parameters in 1. and 2., and thus identify two ways in which prosecutors are exercising discretion.

6.2 Results

This subsection presents results using the methods described in the previous subsection. To begin, I note that when comparing similarly situated Black and White men using the same methods in section 5.2 I find Black men are more likely to have a partial (e.g., charge within a case) or full case dismissal. Specifically, I find Black individuals are 2.8 percentage points (10%) more likely to receive a full case dismissal and 2.7 percentage points (10%) more likely to receive a partial dismissal than a similarly situated White individual.[48]

Then, I turn to my empirical results that leverage variation that generates more causal estimates. Here I start by presenting Figures 5a and 5b, which show the 2SLS and OLS point estimates of the effect non-prosecution has on the probability that a Black or White defendant is re-arrested 3, 6, 9, 12, and 18 months after disposition, respectively. A few prominent patterns deserve attention. I include sets of estimates that include controls and estimates that do not.

First, in all cases, the 2SLS estimates are significantly larger in magnitude than the OLS estimates. The substantial differences between the OLS and 2SLS estimates may be driven by selection bias and/or heterogeneity in the treatment effect of dismissal. The former can be thought of as differences driven by unobservables that drive prosecutorial decision-making and are correlated with the outcome and thus bias OLS, whereas the latter can be thought of as differences in the populations that drive the LATE identified by 2SLS and the ATE identified by OLS. In other words, even in the absence of concerns of endogeneity—thus allowing OLS to provide a completely unbiased estimate of the causal effect of dismissals— OLS and 2SLS could produce very different estimates. In this scenario, the difference would be entirely driven by the fact that 2SLS identifies the ATE for the subpopulation of compliers, while OLS identifies the ATE of the entire population. The tables in Appendix D replicate

---

[48]These results are both significant at the 99% level after clustering standard errors by defendant-prosecutor.

the results in Figures 5-6, with the addition of a third column for each time step that represents the estimates produced by the weighted OLS, where the weights are the LATE weights derived in section D. The results in Appendix D show re-weighted estimates fall between the smaller OLS estimates and larger 2SLS estimates, being much closer to the former than the latter. This finding implies some heterogeneity exists in the effect of non-prosecution, due to differences in the complier population, but the substantial gap between the two models is likely driven in large part by bias introduced by unobservables that correlate with prosecutorial decisions and future contact with the criminal legal system.

Second, in all cases, dismissal has a monotonically larger effect as the time horizon lengthens. This finding can partially be explained by the incapacitation effect, whereby marginally prosecuted defendants are likely to be incarcerated (70% of the time in the sample) and therefore are less able to re-offend than defendants with marginally non-prosecuted charges. Furthermore, the average carceral sentence for defendants in the sample is approximately nine months, which generally aligns with the jump in my estimates. In addition, this trend can also reflect how prosecution changes the trajectory of an individual's life, causing marginally non-prosecuted defendants to desist from future criminal activity, whereas marginally prosecuted individuals are drawn further in. In summary, I take these estimates as evidence that prosecution has a criminogenic effect.

Finally, I turn to a discussion of the race-specific patterns in my causal estimates, which are the most pertinent results with respect to understanding how the prosecutor's discretion interacts with race. Taking the 18-month estimates as an example, the effect of charge dismissal on the probability of future re-arrest for Black men is significantly larger in magnitude ($\sim -48\%$) than it is for White men ($\sim -14\%$). This pattern is slightly less pronounced when looking at the effect of dismissals on future re-conviction, with estimates of $\sim -45\%$ and $\sim -20\%$ for Black and White individuals, respectively. These results imply that if prosecutors reduce prosecution, future criminal activity for both Black and White
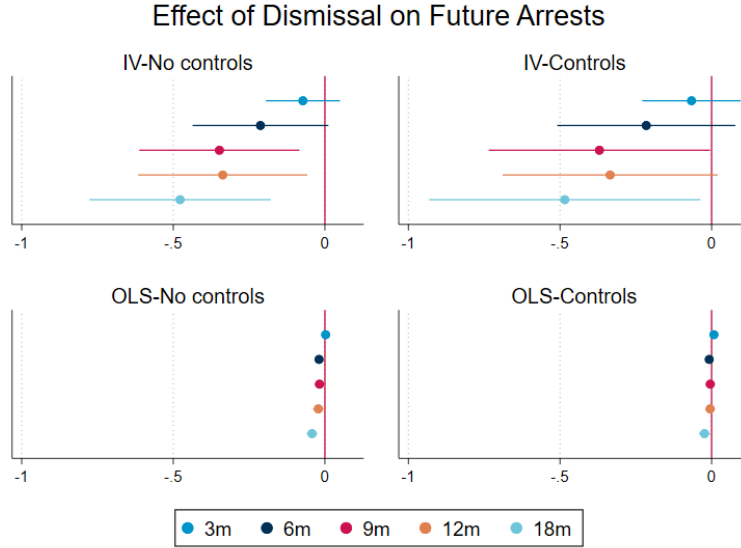
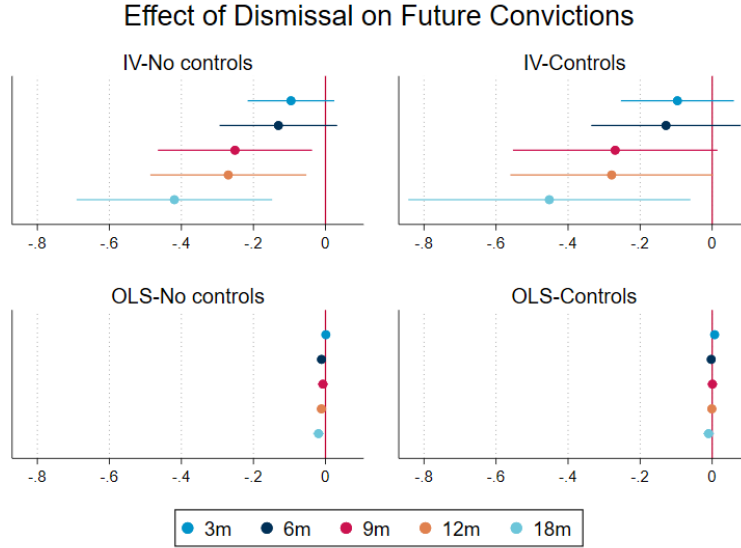**Effect of Dismissal on Future Arrests**

(a) Black subsample

Sample size is 11,129 charges. The mean re-arrest rate after 3, 6, 9, 12, and 18 months in the entire sample of prosecuted Black defendants is 6.4%, 15%, 19%, 24%, and 32%, respectively. The Kleibergen-Paap statistic is 18.11. Errors are two-way clustered at the defendant-prosecutor level.

Effect of Dismissal on Future Arrests
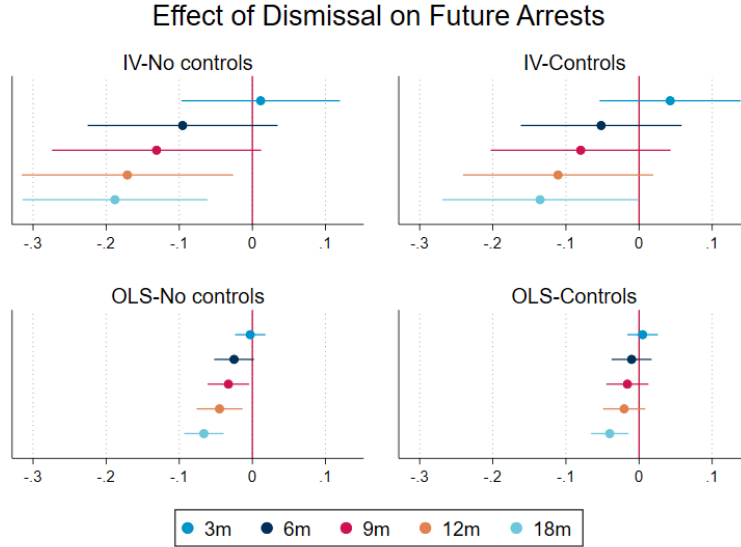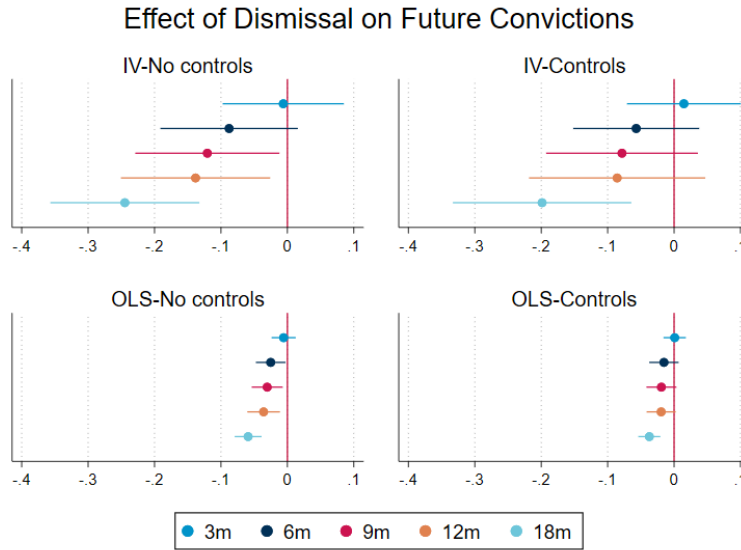
(b) White subsample

Sample size is 7,262 charges. The mean re-arrest rate after 3, 6, 9, 12, and 18 months in the entire sample of prosecuted White defendants is 6.7%, 13%, 18%, 22%, and 29%, respectively. The Kleibergen-Paap statistic is 17.36. Errors are two-way clustered at the defendant-prosecutor level.

Figure 5: Estimates from OLS and 2SLS with and without controls of the effect dismissal has on the probability of future re-arrest 3-18 months after disposition for (a) Black and (b) White male defendants.

Effect of Dismissal on Future Convictions

(a) Black subsample

Sample size is 11,129 charges. The mean re-arrest rate after 3, 6, 9, 12, and 18 months in the entire sample of prosecuted Black defendants is 4.3%, 9.4%, 12%, 15%, and 19%, respectively. The Kleibergen-Paap statistic is 18.11. Errors are two-way clustered at the defendant-prosecutor level.

Effect of Dismissal on Future Convictions

(b) White subsample

Sample size is 7,262 charges. The mean re-arrest rate after 3, 6, 9, 12, and 18 months in the entire sample of prosecuted White defendants is 5.0%, 10%, 13%, 16%, and 21%, respectively. The Kleibergen-Paap statistic is 17.36. Errors are two-way clustered at the defendant-prosecutor level.

Figure 6: Estimates from OLS and 2SLS with and without controls of the effect dismissal has on the probability of future re-conviction 3-18 months after disposition for (a) Black and (b) White male defendants

individuals would *reduce*, but more so for the former.

To better understand what mechanisms are at play, I look to insights from my theoretical model developed in section 6.1.2. This model allows me to jointly use information from estimates on re-arrests and re-conviction to infer how prosecutors may be using additional convictions and/or if they are setting different requirements for dismissals by race. To remind the reader, the former is defined as

$$\Delta\theta^B - \Delta\theta^W = \left[\frac{\tilde{\pi}^B}{\tilde{\alpha}^B} - \frac{\pi^B}{\alpha^B}\right] - \left[\frac{\tilde{\pi}^W}{\tilde{\alpha}^W} - \frac{\pi^W}{\alpha^W}\right], \tag{12}$$

where π^R ($\tilde{\pi}^R$) represents the probability of conviction and α^R ($\tilde{\alpha}^R$) represents the probability of arrest for dismissed (prosecuted) compliers of race R. And the latter is defined as

$$\theta^B - \theta^W = \frac{\pi^B}{\alpha^B} - \frac{\pi^W}{\alpha^W} \tag{13}$$

or

$$\tilde{\theta}^B - \tilde{\theta}^W = \frac{\tilde{\pi}^B}{\tilde{\alpha}^B} - \frac{\tilde{\pi}^W}{\tilde{\alpha}^W}. \tag{14}$$

Figures 5-6 represent estimates of $\hat{\tilde{\pi}} - \hat{\pi}$ and $\hat{\tilde{\alpha}} - \hat{\alpha}$. Therefore, I use the methods detailed in Appendix E to further decompose these results to obtain estimates of all parameters in equations 12, 13, and 14. Table 4 presents the estimates for each parameter using bootstrap methods to obtain standard errors. These parameters can then be plugged in to get estimates of equations 12, 13, and 14.

My estimates indicate $\Delta\theta^B - \Delta\theta^W < 0$, which implies prosecutors penalize Black defendants less than White defendants for additional convictions when making prosecution decisions. This result is in agreement with my previous results in which Black defendants were penalized less than similarly situated White defendants for prior convictions on the intensive margin (e.g., carceral sentence lengths, Figures 2a-2c). To reiterate, I assume ad-

Table 4: Model Parameter Estimates

Object	White	Black	Difference
Δ_A	0.12	0.49	0.37
	(0.09)	(0.17)	(0.2)
Δ_C	0.19	0.46	0.28
	(0.07)	(0.15)	(0.18)
$\tilde{\alpha}$	0.31	0.63	0.32
	(0.07)	(0.14)	(0.16)
α	0.19	0.14	-0.05
	(0.06)	(0.09)	(0.11)
$\tilde{\pi}$	0.23	0.56	0.32
	(0.06)	(0.13)	(0.15)
π	0.04	0.09	0.05
	(0.04)	(0.07)	(0.09)

Note: All estimates obtained via bootstrapping using 1,000 iterations, with stratified sampling at the fixed effect level. These estimates are of the future probabilities of re-arrest (α) and re-conviction (π) for prosecuted (denoted with tilde) and dismissed compliers 18 months after disposition.

ditional convictions do not vary the future case characteristics/conduct of individuals in a race-specific manner. If this assumption fails, my result $\Delta\theta^B - \Delta\theta^W < 0$ could be driven by additional convictions causing an increase in conduct and / or case characteristics more for White individuals than for Black individuals.[49] If the reverse is true, my estimates are actually a lower bound on the prior-conviction discount given to Black individuals.

Moving on to the second measure of prosecutorial discretion, my estimates in Table 4 imply $\theta^B - \theta^W > 0$ and $\tilde{\theta}^B - \tilde{\theta}^W > 0$, consistent with prosecutors setting more stringent requirements for non-prosecution of charges brought against Black men relative to White men. However, the assumptions required to make this claim are stronger than those in the previous paragraph, because we are comparing marginal individuals *across* race. For the positive gap between the probability of future conviction conditional on future arrest for Black ($\theta^B/\tilde{\theta}^B$) and White ($\theta^W/\tilde{\theta}^W$) marginally non-prosecuted/prosecuted individuals to be entirely driven by differential thresholds, I require that criminal records and conduct be identical across both groups. Table 2 shows Black individuals are more likely to have longer

---

[49]By "increasing", I mean conduct or case characteristics that increase the probability of conviction.

criminal records, suggesting this assumption may fail, specifically in the direction that would drive the positive estimate (e.g., Black defendants are more likely to have longer criminal records, and criminal records weakly increase the probability of conviction). Therefore, I look to variation from a pilot screening program to offer suggestive evidence that the imbalance over prior convictions does not explain the differences $\theta^B - \theta^W$ or $\tilde{\theta}^B - \tilde{\theta}^W$.

As detailed in Appendix G, the jurisdiction of interest implemented a pilot screening program, in which a retired prosecutor would pre-screen cases before sending them to the office for assignment. The motivation was to apply consistent guidelines to remove cases that would eventually be dismissed based on historic office practices. Furthermore, I conducted a survey that the screener completed after each case was screened and conducted multiple interviews with the screener. Of interest, a question allowed the screener to indicate whether they believed the office should consider dismissing similar cases in the future and provide reasoning. Qualitative analysis of open-ended reasoning and discussions with the screener suggest their threshold for prosecution is likely race neutral and determined by case facts and prior criminal conduct. Therefore, I argue that any differences in non-prosecution rates of charges eligible for screening before and after the program can be interpreted as driven by the differences between the status quo and race-neutral thresholds. In other words, if the prosecutors set race-neutral thresholds, I would expect the change in non-prosecution rates caused by the screener to not exhibit racial heterogeneity.[50] As illustrated in Figure 17, the effect of screening exhibits striking racial heterogeneity, with the non-prosecution rate substantially increasing for Black men and insignificantly decreasing for White men. These results suggest that—at least for the subset of low-level charges eligible—prosecutors set more stringent requirements for non-prosecution of cases involving Black men.[51] I use these results to corroborate my previous findings that suggest prosecutors pursue marginal

---

[50]I make no claim about whether one should expect the level of non-prosecution to remain constant, because the screener may simply set uniformly higher (or lower) requirements for non-prosecution.

[51]If instead of being race neutral the screener is actually biased against White defendants, this fact could also explain the results and allows for the possibility that historical decisions used race-neutral thresholds.

charges that involve less severe conduct for Black individuals than for White individuals.

In summary, on the extensive margin, I find mixed evidence in regards to how prosecutors' discretion affects racial disparities. To start, my results suggest prosecutors utilize additional convictions to increase the probability of conviction less for Black men than for White men. These results are consistent with the intensive-margin results. In addition, my findings suggest prosecutors dismiss charges at higher rates for Black individuals on average; however, this dismissal rate is not enough to equalize marginal conduct, suggesting some racial bias against Black individuals. This bias may be driven in part by institutional constraints paired with the composition of charges that are selected by law enforcement. In the next section, I document patterns in the selection process that draws individuals into the office as a potential reason for why charges against Black defendants are dismissed at higher rates, yet the marginal charge for Black defendants seems to be for lesser conduct/behavior.

7 Policing

To summarize, so far, I have shown similarly situated Black defendants receive shorter sentences than comparable White defendants. And at the extensive margin, I find charges against Black defendants are dismissed at higher rates than for comparable White defendants, yet the marginal charge for Black defendants seems to be for lesser conduct/behavior. These two results point toward the pool of Black defendants entering the office as having lower levels of underlying "criminality" than white arrestees. In this section, I briefly examine patterns in the selection process (i.e., policing) that brings charges into the office.

Black people make up a disproportionate share of arrests compared with their share of the general population. As shown in Figure 7a, the ratio of Black to White arrests per capita is stable over time, hovering around five. Therefore, Black men in the jurisdiction are five times more likely to be arrested. Figure 7b shows the disparity in arrests is greater for

younger people, with the ratio rising to 6.5 for people ages 15-24.

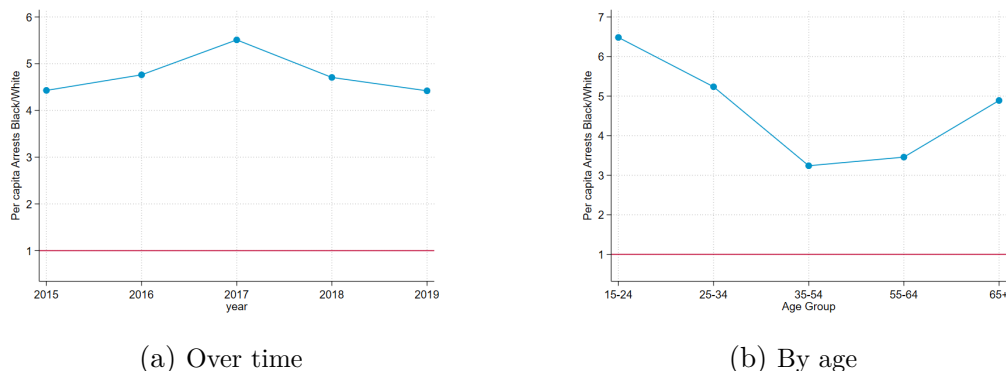(a) Over time　　　　　　　　　　　　　(b) By age

Figure 7: Ratio of Black arrests per capita to White arrests per capita. The red line represents equal representation by race.

Theoretically, these disparities can be driven by two mechanisms; either Black individuals are far more likely to engage in criminal activity and/or Black individuals are much more likely to be in contact with police. To provide empirical evidence that speaks to the former assumption, I focus on patterns in drug arrests, because I have reliable proxies for underlying drug use and sales. My approach is similar to Lum and Isaac (2016).

First, I use a highly predictive model of drug use and apply it to a synthetic population provided by Research Triangle Institute (RTI) International (Wheaton, 2014). The synthetic population was created by placing households—and the individuals within them—from the publicly available micro-data from the 2010 Decennial Census into block groups according to the marginal distributions of the age of the head of household, household income, household size, and race of heads of households.[52] The resulting data allow me to create an average propensity to use or sell drugs by census block group. Before using the geographic information contained in these data, I first compare the overall prevalence of drug use and sales in the population. I find Black individuals are around 10% less likely to have recently used illicit non-marijuana drugs.[53] One may worry the portfolios of drug types are different by race. In

---

[52]More details on the synthetic population are provided in Appendix H.

[53]This finding is also reflected in the raw response rates contained in Table 3, where 9.8% of White individuals report using non-marijuana drugs recently, versus 7% of Black individuals.

46

this case, the arrest disparities may be driven by Black individuals using "more dangerous" drugs (e.g., drugs that are more detrimental to public health). However, data derived from death certificates show White people are nearly twice as likely to die from a drug overdoses in the jurisdiction, making the possibility that drug type is driving the disparity unlikely.

Moving on to policing patterns, I geolocate each arrestee's residential address for arrests involving drugs to get the average number of drug arrests of people residing within a block group. I also find five-year estimates of the percentage of residents who are Black by block group from the American Community Survey. Figures 8 and 9 illustrate similar trends, with drug arrestees coming from predominantly Black neighborhoods, regardless of predicted drug use or drug sales.[54] These findings mirror results found in Lum and Isaac (2016) and suggest police strategies can expose Black residents to a higher risk of arrest, given similar levels of conduct.

However, alternative explanations exist, and although Figures 8 and 9 are visually striking, they do not account for any potentially important contextual variables. In addition, the data have a spatial dimension that, if not accounted for, may bias any estimated correlations. Unfortunately, due to privacy concerns, I am unable to provide a visualization of the spatial patterns of arrests per capita; however, I perform a global Moran's I test on both drug arrests per capita and the residuals after accounting for a set of controls. The intuition behind this test is to see if a block group's value can be predicted by the mean value of neighboring block groups.[55] This index can vary from 1 (strong clustering) to -1 (strong dispersion). In my data, the global Moran's I statistic is 0.58 and 0.36 for drug arrests per capita and the residuals after controlling for propensity to use/sell drugs, foot traffic, and median income. Both statistics are large and statistically significant, suggesting the need to account for spatial auto-correlation. Therefore, to account for additional controls and

---

[54]Interestingly, some positive correlation between predicted drug sales and drug arrests seems to exist.
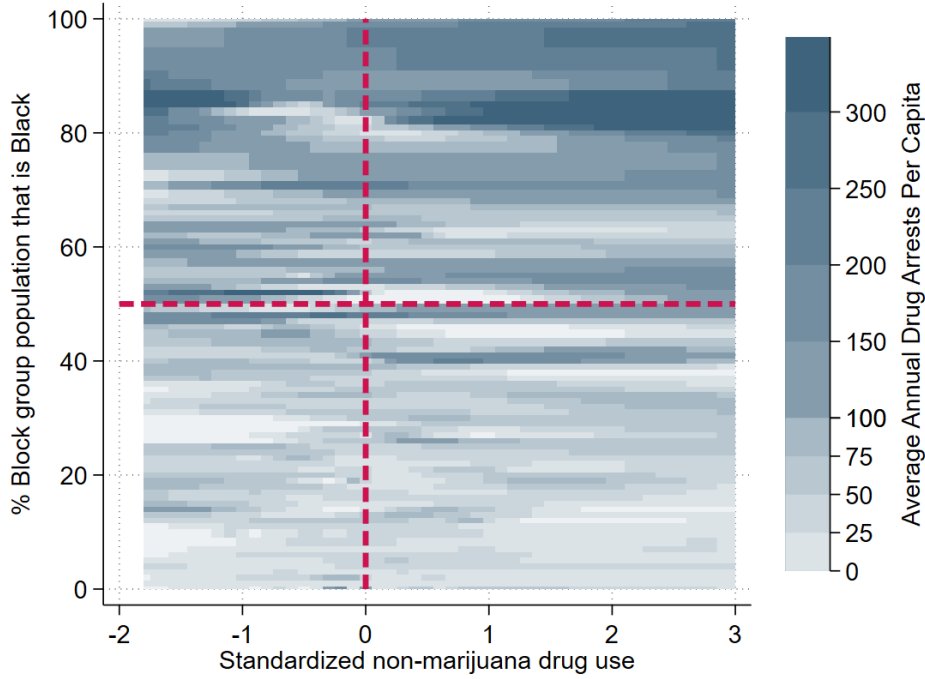[55]Neighbors are defined as block groups that share a border.

Figure 8: Contour plot showing the relationship between block group characteristics—percentage Black and proxies for drug use—and the average number of drug arrests per 10,000 residents per year. The vertical red line denotes the average drug use, and the horizontal red line represents block groups with an even number of Black and White residents.

respect the spatial structure of the data, I estimate a spatial auto-regressive (SAR) model:

$$y_i = \alpha + \beta B_i + \boldsymbol{\theta} X_i + \boldsymbol{\gamma} W y + \epsilon_i , \tag{15}$$

where y_i, B_i, and $\boldsymbol{X}_i$ are the annual number of drug arrestees who reside in, percentage of residents who are Black, and a vector of controls for each block group i. I also include matrices $\boldsymbol{W}$ and $\boldsymbol{y}$, which, when multiplied, simply add spatial lags of block group i's "neighbors" drug arrests per capita to the model.[56] The choice of a spatial weighting matrix defines what is meant by "neighbors". For the results that follow, I use the first-order contiguous neighbors (e.g., block groups that share a border) and row normalization.[57] Next,

---

[56]Alternatively, I can control for a spatial lag of the error term. My results do not change; however, the R^2 decreases, which indicates the lagged-dependant-variable model is preferred.

[57]My results are not substantively different if I include second-order neighbors (i.e., neighbors of neighbors) in the model.
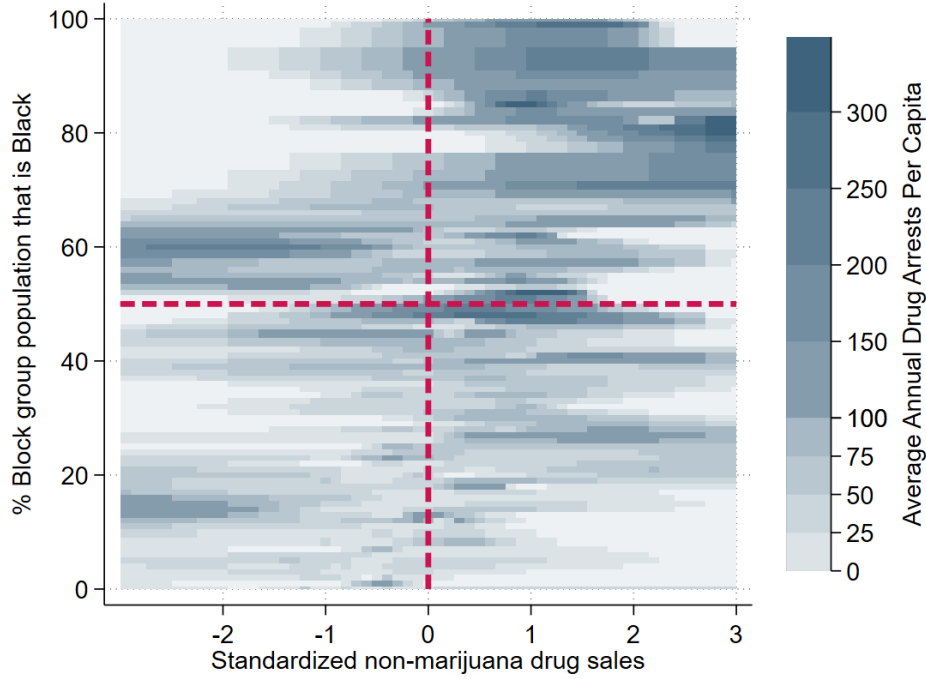
Figure 9: Contour plot showing the relationship between block group characteristics—percentage Black and proxies for drug sales—and the average number of drug arrests per 10,000 residents per year. The vertical red line denotes the average drug sales, and the horizontal red line represents block groups with an even number of Black and White residents.

I include some basic controls that are likely to change the nature and/or propensity to use or sell drugs. To begin, as discussed in length in Mitchell and Caudy (2015), the *nature* of drug use and/or sales may differ by race. For example, Black drug users (and sellers) may use/sell drugs in more public areas that make detection by police more likely. To proxy for the "publicness" of a block group, I use high-quality foot traffic data derived from SafeGraph cell phone data (see Appendix J for details). To better control the general level of income in an area, I also include the median income. Therefore, the estimate of β represents the correlation between what percentage of a block group's residents are Black and the average drug arrests per capita after accounting for foot traffic, median income, predicted drug use and drug sales (non-linear functions of a constellation of demographics), and the number of drug arrests per capita in the surrounding area. The estimate of $\beta = 0.33$ is significant at the 99% level (robust standard error=0.10, p-value=0.001). Figure 10 visualizes the effect

of the racial composition of a block group on the number of annual drug arrests per 10,000 residents holding foot traffic, median income, predicted drug use, predicted drug sales, and drug arrest patterns of neighboring block groups.
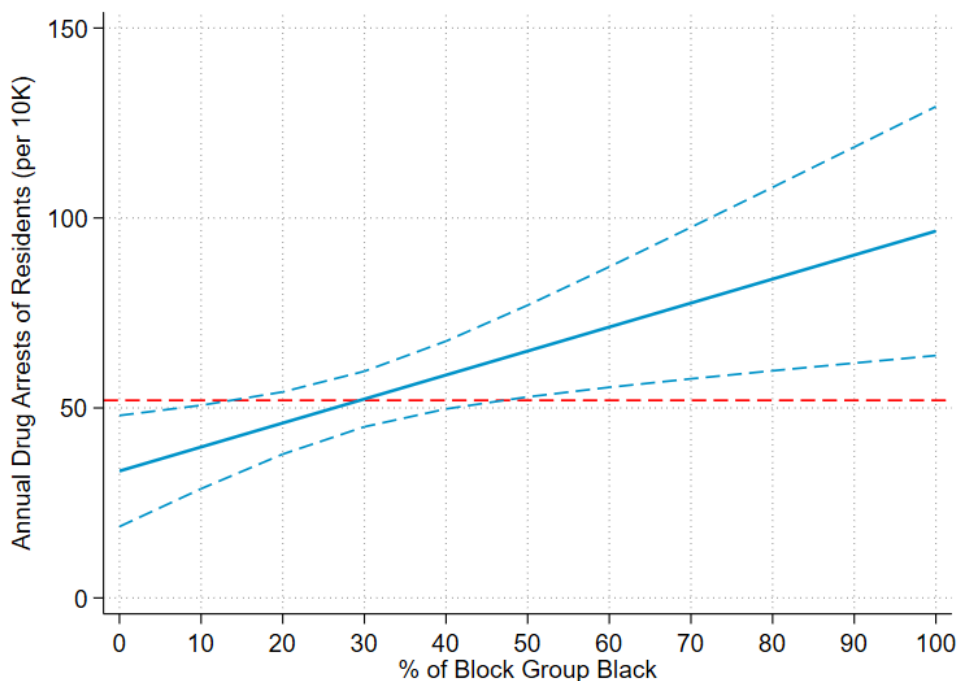

Figure 10: Plot of predicted annual drug arrests of residents of a block group (per 10,000) as a function of the percentage of the population in a block group that is Black. Horizontal dashed red line represents the mean of all block groups.

I take this result as suggestive evidence that Black residents are more likely to be arrested, given similar levels of conduct, thus providing a potential explanation for why the marginal Black defendant has a lower-quality case/underlying criminal propensity than his White counterpart. I want to stress that these results are correlational in nature, and more work would need to be done to uncover the types of bias that drive the total arrest disparities that I documented. Moreover, these results do not imply police are necessarily engaging in direct bias, because we saw with respect to incarceration that these disparities may be driven by systemic features. For example, imagine a likely policing strategy that simply conditions on the average number of drugs found in the prior time period; this strategy could unknowingly reinforce policies set in the past that were driven by racial animus. Because

drug use is fairly uniform by race (i.e., police are likely to "find drugs wherever they look"), the cycle will continually reinforce itself.[58]

8 Discussion

This paper uses high-quality administrative data from a local prosecutor's office to study prosecutorial decision-making. I find that after adjusting for a rich set of case and defendant characteristics, Black defendants receive significantly shorter sentences on average—suggesting prosecutors *reduce* total disparities. This difference between unconditional and conditional disparity in carceral length indicates all of the total disparity is driven by systemic—or upstream—features that factor into a prosecutor's decision rule. Using a novel decomposition method from Gelbach (2016), I find that, out of all observable features, prior criminal convictions drive the majority of this systemic bias. Furthermore, I find that, on average, prior convictions increase carceral sentences; however, they increase much less for Black individuals than for similar White individuals, suggesting race influences how prosecutors interpret signals from upstream sources.

While interesting in themselves, the results regarding incarceration are descriptive in nature. On the extensive margin—where prosecutors enjoy almost unchecked power—I leverage quasi-random assignment of cases to prosecutors to identify causal estimates of how prosecution influeces future re-arrest and re-conviction. Using a simple theoretical model, I show prosecutors use additional prior convictions to increase the probability of conviction less for marginal Black individuals than for White individuals at the margin, echoing the results regarding incarceration. This finding suggests policies aimed at removing information from prosecutors—specifically, information regarding the race of the individual—may inadvertently exacerbate existing racial disparities. The reason is that without being able to

---

[58]See a similar example in Lum and Isaac (2016).

observe race, prosecutors may be unable to use their discretion to discount signals that they believe may be contaminated by racial bias, namely, prior criminal records. This mechanism relates to a large recent literature on fairness in algorithms.[59] For example, in their work on college admissions, Kleinberg et al. (2018) compare algorithms that are allowed to use race as a predictor versus those that are "blinded" to race. The authors find race-blind algorithms *decrease* fairness in access to higher education. The mechanism is similar to the one hypothesized in this paper, whereby the race-aware algorithms are able to learn race-specific relationships between predictors and outcomes that promote fairness (e.g., high school test scores may have less predictive power over future college success for Black students).

My results have clear policy implications. First, as alluded to earlier, a recent push has been made to "color blind" prosecutors (Sah et al., 2015; Chohlas-Wood et al., 2021). In fact, California recently passed a bill that mandates all prosecutor's offices in the state implement "race-blind charging" by 2025. The logic behind these proposals is that without the ability to condition on race, prosecutors are unable to compound racial disparities. However, as the results in this paper and Shaffer (2022) show, these policies effectively "lock in" racial bias baked into other signals, namely, criminal histories.[60] Similarly, my results suggest removing discretion by forcing similarly situated defendants to receive identical sentences may worsen existing racial disparities. More specifically, my principal results on the intensive margin suggest forcing defendants' criminal histories to map in a race-neutral way into incarceration would worsen the racial gap in average carceral lengths.

I also find that, on average, Black individuals are more likely to have their cases/charges dismissed than comparable White individuals. However, I find evidence the *marginal* charge brought against Black individuals to be of lower quality, perhaps due to the selection process that brings charges into the office and institutional policies. An extreme example would

---

[59]See Mitchell et al. (2021) for a recent review of more examples and definitions.

[60]However, blinding other case features, such as victim race, may mitigate disparities. Therefore, I echo Chohlas-Wood et al. (2021)'s argument that more randomized trials are required to determine the true costs and benefits of technologies that remove race from prosecutors' information set.

be a scenario where Black individuals were only arrested for low-level possession of drugs while White individuals were exclusively arrested for more serious violent crimes. Given the absence of an office policy to dismiss all low-level possession cases, some of the cases against Black defendants will be prosecuted. In this world, Black defendants would likely see significantly higher dismissal rates. However, the marginal charge for Black individuals would be for lower conduct than for marginal White individuals.[61] I find evidence suggestive of the marginal conduct being relatively lower for Black defendants than for White defendants, which implies prosecutors pursue lower-level charges against Black individuals. These results suggest policies aimed at reducing the number of low-level and/or low-quality charges that are prosecuted may benefit Black individuals more.

Finally, and likely of greatest concern to policymakers, is my results imply not prosecuting the marginal non-violent charge does not lead to more future arrests or convictions for a defendant or harm public safety. Rather, not prosecuting such low-level offenses greatly reduces the probability of future arrests and convictions for both Black and White defendants, with a larger effect for the former. This result is similar to those found in Mueller-Smith and Schnepel (2021) and Agan et al. (2021) that suggest policies aimed at reducing contact with the criminal legal system may actually decrease future criminal activity and thus improve public safety. Such results call for more research into how the current system of arrest, prosecution, and punishment actually relate to crime and public safety.

---

[61]I stress this example is strictly illustrative and cannot be exactly what drives my results, because my sample does not include violent crimes.

References

Abadie, Alberto, "Bootstrap tests for distributional treatment effects in instrumental variable models," *Journal of the American Statistical Association*, 2002, *97* (457), 284–292.

Agan, Amanda and Sonja Starr, "Ban the box, criminal records, and racial discrimination: A field experiment," *Quarterly Journal of Economics*, 2018, *133* (1), 191–235.

_ , **Jennifer L Doleac, and Anna Harvey**, "Prosecutorial Reform and Local Crime Rates," 2022.

Agan, Amanda Y., Jennifer L. Doleac, and Anna Harvey, "Misdemeanor Prosecution," 2021.

American Civil Liberties Union, "Prosecutorial Reform," 2022.

Andresen, Martin Eckhoff, "Exploring marginal treatment effects: Flexible estimation using Stata," *Stata Journal*, 2018, *18* (1), 118–158.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters, "Explaining charter school effectiveness," *American Economic Journal: Applied Economics*, 2013, *5* (4), 1–27.

Angrist, Joshua, Peter Hull, and Christopher Walters, "Methods for Measuring School Effectiveness," 2022, pp. 1–67.

Anwar, Shamena and Hanming Fang, "An alternative test of racial prejudice in motor vehicle searches: Theory and evidence," *American Economic Review*, 2006, *96* (1), 127–151.

Arnold, David, Will Dobbie, and Crystal S. Yang, "Racial bias in bail decisions," *Quarterly Journal of Economics*, 2018, *133* (4), 1885–1932.

Bates, Savannah, Vasiliy Leonenko, James Rineer, and Georgiy Bobashev, "Using synthetic populations to understand geospatial patterns in opioid related overdose and predicted opioid misuse," *Computational and Mathematical Organization Theory*, 2019, *25* (1), 36–47.

Batista, Gustavo E A P A, Ana L C Bazzan, and Maria Carolina Monard, "Balancing Training Data for Automated Annotation of Keywords: a Case Study," *In Proceedings of the Second Brazilian Workshop on Bioinformatics*, 2003, (January), 35–43.

Beckett, Katherine, Kris Nyrop, and Lori Pfingst, "Race, drugs, and policing: Understanding disparities in drug delivery arrests," *Criminology*, 2006, *44* (1), 105–137.

_ , _ , _ , **and Melissa Bowen**, "Drug use, drug possession arrests, and the question of race: Lessons from Seattle," *Social Problems*, 2005, *52* (3), 419–441.

Bhuller, Manudeep, Gordon Dahl, Katrine Løken, and Magne Mogstad, "Incarceration, Recidivism and Employment," *Journal of Political Economy*, 2020, *128* (4), 1269–1324.

Bibas, Stephanos, "Prosecutorial Regulation versus Prosecutorial Accountability," *University of Pennsylvania Law Review*, 2009, *157* (4), 959–1016.

Blinder, Alan S., "Wage Discrimination : Reduced Form and Structural Estimates," *The Journal of Human Resources*, 1973, *8* (4), 436–455.

Bohren, J. Aislinn, Peter Hull, and Alex Imas, "Systemic Discrimination: Theory and Measurement," 2022.

Brame, Robert, Shawn D. Bushway, Ray Paternoster, and Michael G. Turner, "Demographic Patterns of Cumulative Arrest Prevalence By Ages 18 and 23," *Crime &*

Delinquency, 2014, *60* (3), 471–486.

Carson, E .Ann, "Prisoners in 2019," Technical Report October 2020.

Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, 2002, *16*, 321–357.

Chohlas-Wood, Alex, Joe Nudell, Keniel Yao, Zhiyuan Jerry Lin, Julian Nyarko, and Sharad Goel, "Blind Justice: Algorithmically Masking Race in Charging Decisions," *AIES 2021 - Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 2021, *499* (1954), 35–45.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg, "From LATE to MTE: Alternative methods for the evaluation of policy interventions," *Labour Economics*, 2016, *41*, 47–60.

Dobbie, Will, Jacob Goldin, and Crystal S. Yang, "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, 2018, *108* (2), 201–240.

Doleac, Jennifer L, "Racial Bias in the Criminal Justice System," in Paolo Buonanno, Juan Vargas, and Paolo Vanin, eds., *A Modern Guide to Economics of Crime*, Edward Elgar Publishing, 2022.

Doleac, Jennifer L. and Benjamin Hansen, "The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden," *Journal of Labor Economics*, 2020, *38* (2), 321–374.

Dunlea, R. R., ""No idea whether he's Black, White, or purple": Colorblindness and cultural scripting in prosecution," *Criminology*, 2021, *60* (2), 237–262.

FBI Uniform Crime Report, "Crime in the United States," Technical Report 2019.

Fendrich, Michael and Timothy P. Johnson, "Race/ethnicity differences in the validity of self-reported drug use: Results from a household survey," *Journal of Urban Health*, 2005, *82* (3), 67–81.

_ , _ , **Joseph S. Wislar, Amy Hubbell, and Vina Spiehler**, "The utility of drug testing in epidemiological research: Results from a general population survey," *Addiction*, 2004, *99* (2), 197–208.

Frandsen, Brigham, "TESTJFE: Stata module to perform test for instrument validity in the judge fixed effects design," 2020.

Frandsen, Brigham R., Lars John Lefgren, and Emily C. Leslie, "Judging Judge Fixed Effects," 2019.

Gelbach, Jonah B., "When do covariates matter? And which ones, and how much?," *Journal of Labor Economics*, 2016, *34* (2), 509–543.

Goncalves, Felipe and Steven Mello, "Does the Punishment Fit the Crime? Speeding Fines and Recidivism," *SSRN Electronic Journal*, 2017.

Greiner, D. James and Donald B. Rubin, "Causal effects of perceived immutable characteristics," *Review of Economics and Statistics*, 2011, *93* (3), 775–785.

Harrington, Emma and Hannah Shaffer, "Brokers of Bias in the Criminal System: Do Prosecutors Compound or Attenuate Earlier Racial Disparities?," 2021.

Heckman, James J. and Edward J. Vytlacil, "Local instrumental variables and latent variable models for identifying and bounding treatment effects," *Proceedings of the National Academy of Sciences of the United States of America*, 1999, *96* (8), 4730–4734.

___ **and** ___ , "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments," in "Handbook of Econometrics," Vol. 6 2007, chapter Chapter 71, pp. 4875–5143.

___ **and Edward Vytlacil**, "Policy-Relevant Treatment Effects," *American Economic Review*, 5 2001, *91* (2), 107–111.

___ **and** ___ , *Structural equations, treatment effects, and econometric policy evaluation*, Vol. 73 2005.

Heckman, James J, Sergio Urzua, and Edward Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *The Review of Economics and Statistics*, 2006, *88* (August), 389–432.

Hogan, Thomas P., "De-prosecution and death: A synthetic control analysis ofthe impact ofde-prosecution on homicides," *Criminology & Public Policy*, 2022, pp. 1–46.

Holm, Sture, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, 1979, *6* (2), 65–70.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467.

Johnson, Timothy P. and Phillip J. Bowman, "Cross-cultural sources of measurement error in substance use surveys," *Substance Use and Misuse*, 2003, *38* (10).

Kitagawa, Evelyn M., "Components of a Difference Between Two Rates," *Journal of the American Statistical Association*, 1955, *50* (272), 1168–1194.

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 2006, *133* (1), 97–126.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ashesh Rambachan, "Algorithmic Fairness," *AEA Papers and Proceedings*, 2018, *108*, 22–27.

Knowles, John, Nicola Persico, and Petra Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 2009, *117* (6), 1155–1159.

Ledgerwood, David M., Bruce A. Goldberger, Nathan K. Risk, Collins E. Lewis, and Rumi Kato Price, "Comparison between self-report and hair analysis of illicit drug use in a community sample of middle-aged men," *Addictive Behaviors*, 2008, *33* (9), 1131–1139.

Lum, Kristian and William Isaac, "To predict and serve?," *Significance*, 2016, *13* (5), 14–19.

McCance-Katz, Elinore, Debra Houry, and Francis Collins, "Testimony on Addressing the Opioid Crisis in America: Prevention, Treatment, and Recovery before the Senate Subcommittee," 2017.

Mitchell, Ojmarrh and Michael S. Caudy, "Examining Racial Disparities in Drug Arrests," *Justice Quarterly*, 2015, *32* (2), 288–313.

___ , **Lyndsay N Boggess, Daniela Oramas Mora, and Tracey L Sticco**, "Are progressive chief prosecutors effective in reducing prison use and cumulative racial / ethnic disadvantage ? Evidence from Florida," *Criminology & Public Policy*, 2022, pp. 1–31.

Mitchell, Shira, Eric Potash, Solon Barocas, Alexander D'Amour, and Kristian Lum, "Algorithmic fairness: Choices, assumptions, and definitions," *Annual Review of Statistics and Its Application*, 2021, *8*, 141–163.

Mueller-Smith, Michael and Kevin T. Schnepel, "Diversion in the Criminal Justice

System," *Review of Economic Studies*, 2021, *88* (2), 883–936.

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 1973, *14* (3), 693–709.

Office of the Surgeon General U.S. Department of Health and Human Services (HHS), "Facing Addiction in America: The Surgeon General's Report on Alcohol, Drugs, and Health," Technical Report, Washington, DC 2016.

Olea, José Luis Montiel and Carolin Pflueger, "A Robust Test for Weak Instruments," *Journal of Business and Economic Statistics*, 2013, *31* (3), 358–369.

Perry, Steven W. and Duren Banks, "Prosecutors in State Courts, 2007 - Statistical Tables," Technical Report December, U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics 2011.

Pfaff, John F., *Locked in : the true causes of mass incarceration – and how to achieve real reform*, New York, NY: Basic Books, 2017.

Rehavi, M Marit and Sonja B Starr, "Racial Disparity in Federal Criminal Sentences," *Journal of Political Economy*, 2014, *122* (6), 1320–1354.

Reuter, Peter, Jonathan P. Caulkins, and Greg Midgette, "Heroin use cannot be measured adequately with a general population survey," *Addiction*, 2021, *116* (10), 2600–2609.

Rios, Victor M., *Punished : Policing the Lives of Black and Latino Boys*, New York University Press, 2011.

Robertson, Christopher, Shima Baradaran Baughman, and Megan S. Wright, "Race and Class: A Randomized Experiment with Prosecutors," *Journal of Empirical Legal Studies*, 2019, *16* (4), 807–847.

Rose, Evan K. and Yotam Shem-Tov, "How does incarceration affect reoffending? Estimating the dose-response function," *Journal of Political Economy*, 2021, *129* (12), 3302–3356.

Sah, Sunita, Christopher T. Robertson, and Shima B. Baughman, "Blinding prosecutors to defendants' race: A policy proposal to reduce unconscious bias in the criminal justice system," *Behavioral Science & Policy*, 2015, *1* (2), 69–76.

Shaffer, Hannah, "Prosecutors, Race, and the Criminal Pipeline," *The University of Chicago Law Review*, 2022.

Shannon, Sarah, Christopher Uggen, Melissa Thompson, Jason Schnittker, and Michael Massoglia, "Growth in the U.S. Ex-Felon and Ex-Prisoner Population, 1948 to 2010," 2010.

Sherrard, Ryan, ""Ban the Box" Policies and Criminal Recidivism," *SSRN Electronic Journal*, 2020.

Simoiu, Camelia, Sam Corbett-Davies, and Sharad Goel, "The problem of inframarginality in outcome tests for discrimination," *Annals of Applied Statistics*, 2017, *11* (3), 1193–1216.

Sklansky, Jordan A., "The Nature and Function of Prosecutorial Power," *Journal of Criminal Law and Criminology*, 2016, *106* (3).

Sloan, CarlyWill, "Racial Bias by Prosecutors: Evidence from Random Assignment," 2022.

Small, Mario L. and Devah Pager, "Sociological perspectives on racial discrimination," *Journal of Economic Perspectives*, 2020, *34* (2), 46–97.

Stashko, Allison and Haritz Garro, "Prosecutor Elections and Police Accountability," 2021.

Tomek, Ivan, "Tomek Link: Two Modifications of CNN," *IEEE Trans. Systems, Man and Cybernetics*, 1976, pp. 769–772.

Tuttle, Cody, "Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums," *SSRN Electronic Journal*, 2021.

West, Jeremy, "Racial Bias in Police Investigations," 2018, (October), 1–36.

Wheaton, W.D., "2010 U.S. Synthetic Population Ver. 1," 2014.

Wu, Jawjeong, "Racial/Ethnic Discrimination and Prosecution: A Meta-Analysis," *Criminal Justice and Behavior*, 2016, *43* (4), 437–458.

A   Descriptive Statistics

Table 5 summarizes the raw data used to construct my main sample of interest.

	Charges	% of Total
Race		
Black	51742	56.0
White	38468	41.6
Other	2267	2.5
Gender		
Male	75879	82.1
Female	16598	17.9
Age		
15-17	2199	2.4
18-24	21241	23.0
25-34	34877	37.7
35-44	20217	21.9
45-54	9862	10.7
55-64	4081	4.4
Charge Outcomes		
Plea	46069	64.5
Dismissal	20143	28.2
Diverted	3279	4.6
Other	1561	2.2
Trial	419	0.6
Disposition Time		
Pre-COVID	45230	48.9
Post-COVID	25221	27.3
Pending	22026	23.8
Arrest Agency		
Agency 1	21448	23.2
Agency 5	19107	20.7
Agency 2	16779	18.1
Other	15691	17.0
Agency 3	13405	14.5
Agency 4	6047	6.5
Defense Type		
Public Defender	41940	45.4
Private	39139	42.3
None	11249	12.2
Unknown	149	0.2
Number of Co-defendants		
Single defendant	71303	77.1
2	13857	15.0
3	4137	4.5
4+	3180	3.4
Statutorily Violent		
Non-violent	80023	86.5
Violent	12454	13.5
Number of victims		
Victimless	55681	60.2
1	25255	27.3
2	6263	6.8
4	3200	3.5
3+	2078	2.2

Table 5: Summary table of charge-level dataset before any filters are applied.

B Validity of Using NSDUH to Infer Drug Behavior

First, and perhaps least concerning, is the sampling frame used for the survey. The NSDUH includes not only a representative sample of households, but also sampling from homeless shelters, rooming or boarding houses, dormitories, migratory work camps, and halfway houses. The two excluded groups that can drive underestimation are unsheltered homeless and the incarcerated. If the goal were to get an exact estimate of the level of drug use within a geographic area, these two excluded groups may bias estimates toward zero because both populations likely use and/or sell drugs at higher rates than the rest of the population. However, as stated earlier, my main goal is to measure the racial gap in drug use and sales in the nonincarcerated population. Therefore, to dramatically affect my analysis, the unsheltered homeless population would need to be large, predominately Black.

The second critique is that the survey is a selected sample because non-response rates are non-random. Again, although this issue is likely to drive underestimation of the drug use/sales levels, for this issue to affect the racial gap, non-responsiveness would need to correlate with race. Data on response rates by race shown in Figure 11 show no large differences by race. The mean response rate from 2015-2020 for the Black sample is 71% compared with a mean response rate of 66% for the White sample. Whether I should expect larger rates of drug use/sales in the non-respondents is *a priori* ambiguous (see Reuter et al. (2021) for discussion). To partially explain the racial differences I observe—whereby White respondents tend to use drugs at a higher rate than Black respondents—non-respondents would need to be significantly less likely to use drugs, thus biasing the difference between White and Black drug use upward.

The third critique is that because drug use and sales are rare, the average rate of use/sales year-over-year for small groups can be driven by a very small number of unweighted observations. The example given to illustrate this point in Reuter et al. (2021) is a spike in drug use for older users driven by only two older respondents. To guard against such

issues, I present pooled estimates that utilize five years of data (2015-2019) and I resample the raw data as well (see section I for details on resampling methods). Additionally, I am comparing large groups because my raw unweighted data comprise 45,534 Black respondents and 229,186 White respondents.

The final critique centers around under-reporting of drug use at both the extensive and intensive margins. To begin, I simply code recent drug use and sales and avoid breaking this information down by drug or intensity. My analysis is still susceptible to under-reporting on the extensive margin (e.g., reporting no use when there was actual use). To drive my results, this under-reporting would need to occur at a significantly higher rate for Black respondents. The existing literature shows under-reporting is in fact more likely for Black respondents, offering a potential explanation for my results.[62] I still believe, however, that even if under-reporting were partially driving some of the racial gap I see in drug use/sales—where White men use drugs more frequently than Black men—it is unlikely to completely reverse this trend and thus account for the enormous disparity I observe in arresting behavior—where Black men are arrested much more frequently.

C   Testing LATE Assumptions

As described in section 6, to interpret the 2SLS estimates as LATE, I need the following assumptions to hold:

1. Random assignment of the instrument

2. Relevance

---

[62]See (Fendrich et al., 2004; Fendrich and Johnson, 2005; Ledgerwood et al., 2008) for studies citing correlations with being Black and under-reporting drug use on surveys and Johnson and Bowman (2003) for a review of older literature with more mixed results. Note all studies documenting a positive correlation are based on small samples. Additionally, in Fendrich and Johnson (2005), the correlation disappears after controlling for socioeconomic status.
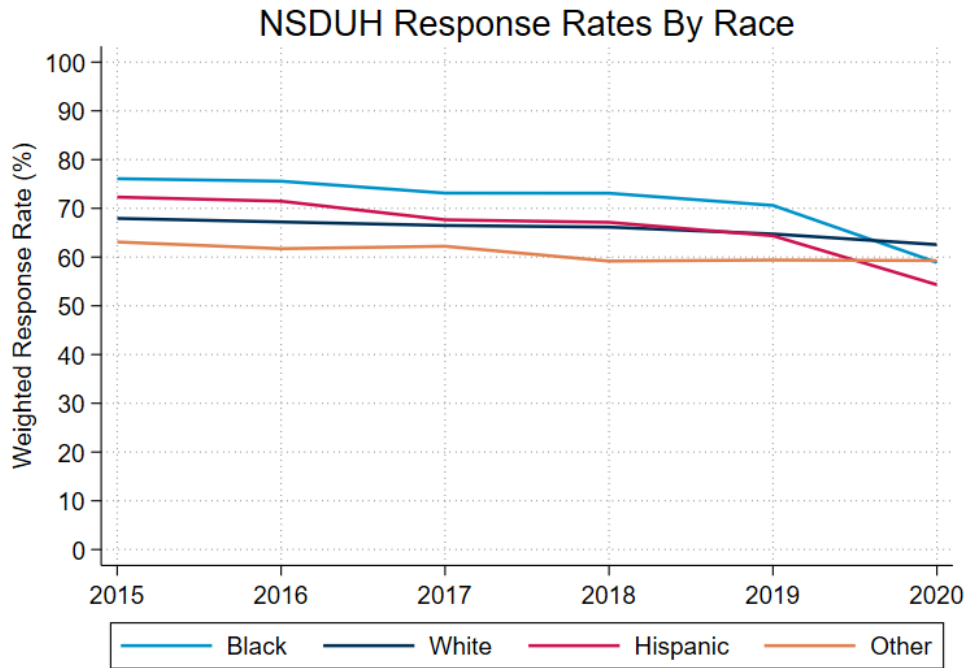
Figure 11: Data compiled from: 2015-2020 Methodological summary and definitions reports accessed here for: 2020, 2019, 2018, 2017, 2016, and 2015

3. Exclusion

4. Monotonicity.

Although only one of these assumptions (relevance) is directly testable, this section will provide suggestive empirical evidence that supports each of them.

C.1 Random Assignment

I begin with random assignment. This assumption holds if charges of the same broad type (i.e., Person, Property, Drugs, Traffic, Guns, and a catch all, "Other") are assigned in a quasi-random manner to prosecutors within a team during a certain time period (i.e., within a three-month "season" and if the arrest occurred on a weekend/weekday). After multiple discussions with the office, I am confident statutorily non-violent charges are assigned in a

quasi-random manner (i.e., first in first out to next person up). To further test this assumption, I residualize (demean) the instrument (equation 5) and a set of observable charge and defendant characteristics with respect to team-crime type-time fixed effects.[63] I then regress the instrument on all observable characteristics and perform a series of linear hypothesis tests that each observable has a statistically significant correlation with the instrument.[64] Additionally, I repeat the same procedure but replace prosecutor leniency with an indicator for whether the charge was dismissed. If charges are truly assigned to prosecutors randomly (after accounting for fixed effects), I should find no observables correlate with prosecutor characteristics—namely, their average propensity to dismiss a case. Conversely, thinking these same observables may have large statistically significant impacts on the probability of receiving a dismissal is reasonable. For example, the number of prior convictions should not determine which prosecutor you are assigned to, but will most certainly have an impact on whether you receive a dismissal. Figure 12 presents the results of these tests as well as the R^2 and F-statistics. In both subsamples, I reject the null hypothesis that all of the regression coefficients are equal to 0 when I regress observables onto the instrument, which could imply some imbalances in the assignment of cases to prosecutors along certain dimensions. I note that in my preferred specification, I add controls for all observables listed and therefore must make the stronger assumption that, conditional on all of these observables, the residual variation is quasi-random. This approach is not ideal; however, I am comforted by three patterns: (1) The addition of these controls does not substantively change my main findings; (2) the observables explain a significantly larger portion of variation in dismissal decisions; and (3) the correlations I find are small in magnitude. I interpret these patterns as suggestive that whatever imbalances may exist are not driving my main findings.

---

[63]Observable characteristics include the age of the defendant at arrest, an indicator for whether the charge is the lead charge in the case, the number of prior convictions, the number of concurrent charges, number of co-defendants, and number of victims

[64]To adjust for the fact I am testing multiple hypothesis, I implement the method found in Holm (1979) by using the *test* post-estimation command in Stata—after running *reg instrument covariates, rob*—with the *mtest(holm)* option.
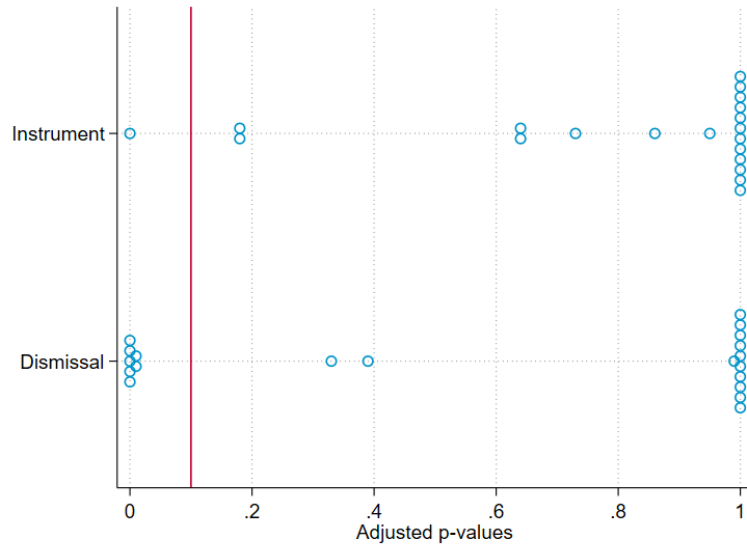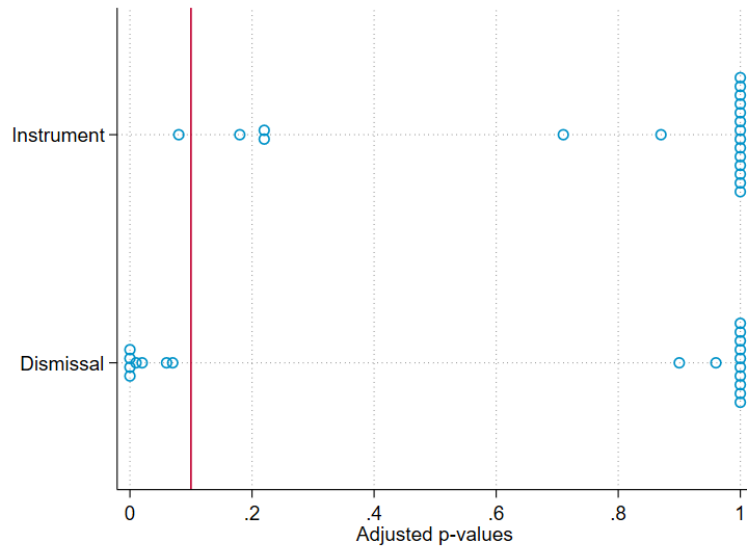
(a) Black subsample
Instrument: $F(20, 11271) = 5.23$ and $R^2 = 0.0098$
Dismissal: $F(20, 11271) = 69.28$ and $R^2 = 0.0940$

(b) White subsample
Instrument: $F(20, 7602) = 8.96$ and $R^2 = 0.0217$
Dismissal: $F(20, 7602) = 50.50$ and $R^2 = 0.0975$

Figure 12: Plot of coefficients that are statistically greater than 0; red line represents p-value=0.10

C.2 Relevance

Figure 13 shows the strong relationship between the instrument—the leave-out propensity to dismiss a case—and the actual probability of a charge not being prosecuted. The clear positive correlation implies being assigned to a more lenient prosecutor does in fact increase the probability of a charge not being prosecuted. Although a clear positive correlation exists, I also confirm the instrument is not "weak". To this point, I present the Kleibergen-Paap F-statistic (Kleibergen and Paap, 2006) in the captions for all figures displaying 2SLS estimates. Additionally, I utilize the *weakivtest* command in Stata that implements the tests outlined in Olea and Pflueger (2013), which also indicates the instrument is not "weak".

C.3 Exclusion

To be confident I am identifying the LATE with my 2SLS estimates requires that prosecutor assignment only affects the future outcomes (e.g., re-arrest, re-conviction, etc.) through the probability of non-prosecution. To start, this exclusion restriction would be violated if the assignment were correlated with other characteristics of the charge and / or the defendant that also affect future outcomes. However, the exclusion restriction could also be violated if assignment were indeed random, but decisions/actions by the prosecutor other than indictment were correlated with the outcome of interest. To further probe this assumption, I follow a similar analysis to that found in section 4.1.3 in Agan et al. (2021) and regress the time it takes a charge to be disposed of on the instrument, covariates, and fixed effects for charges that were prosecuted. I find no statistically significant relationship between the instrument and the time to disposition at the 10% level (p-value=0.113).[65] This result, paired with the results contained in the next subsection, suggest my main results are driven by the decision of the prosecutor whether to prosecute a charge.

---

[65]Because almost all charges that are prosecuted are found guilty (99.81%), I cannot conduct the second test in Agan et al. (2021) on the conviction rate.
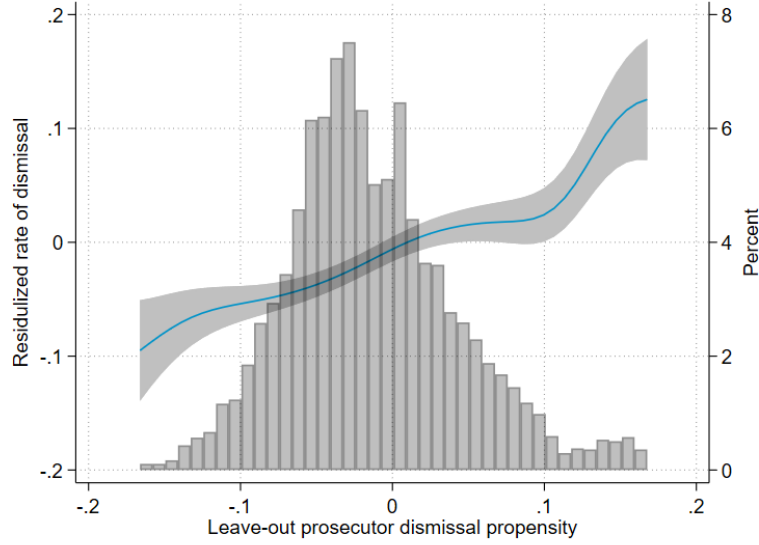
(a) Black sub-sample
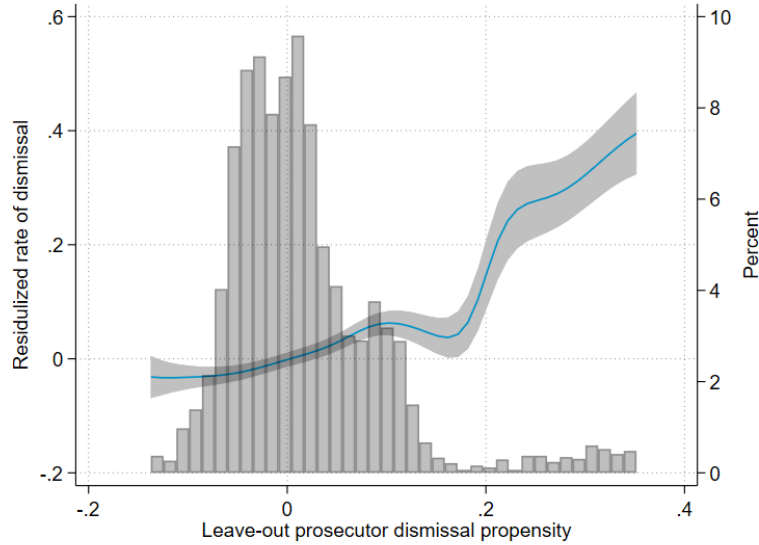
(b) White subsample

Figure 13: Representation of first stage in 2SLS. Both dismissal probability (y-axis) and prosecutor leave-out propensity to dismiss a case (x-axis) are residualized with respect to team by arrest crime type by time of arrest fixed effects.

C.4 Monotonicity

If my causal effects were homogeneous (i.e., if non-prosecution affected all defendants the same regardless of situation), I would only need the prior assumptions in this section to hold

to identify the true causal effect of non-prosecution. However, as evidenced by the heterogeneity by race, this assumption would be a strong one. Fortunately, I can still interpret my estimates as causal if I assume the effect of non-prosecution is monotonic. Montonicity implies that if a defendant's charge was not prosecuted (prosecuted) by prosecutor j, they would also have received a non-prosecution (prosecution) from all prosecutors who were more (less) lenient than j. Testing directly would require observation of an unobservable state—outcomes under a different prosecutor than was assigned—and therefore no direct tests exist. However, recent advances put forth in Frandsen et al. (2019) do allow for a joint test of the monotonicity and exclusion assumptions. I utilize the package $testjfe$ in Stata (Frandsen, 2020) to calculate this statistic, controlling for the covariates and fixed effects used throughout the paper. I fail to reject the null hypothesis that the exclusion and/or monotonicity assumptions are violated. [66]

Furthermore, Frandsen et al. (2019) proves the strict monotonicity assumption can be relaxed to a weaker assumption of "average monotonicity" (see Theorem 5 in Frandsen et al. (2019)). This result can be empirically confirmed by testing if the instrument is relevant over all subsamples in the data.[67] Figure 14 shows the coefficient in front of the leave-out propensity of non-prosecution is, in fact, positive and significant in a multitude of subsamples of the data.[68] I take these two results to suggest monotonicity and exclusion assumptions are reasonable in my setting.

---

[66]The statistic is a weighted average of two statistics with user-chosen weights. Additionally, the user must decide the number of knots to use when fitting the spline-based relationship between the outcome and the dismissal propensity. To confirm the failure to reject was not driven by choice of weights, or the number of knots, I re-ran the test using all pairwise combinations of weights$\in \{0, .25, .5, .75, 1\}$ and knots $\in \{3, 5, 10, 15, 20\}$.

[67]For empirical examples of this test, see (Bhuller et al., 2020; Dobbie et al., 2018; Agan et al., 2021).

[68]Each regression has non-prosecution as the dependant variable and includes the team-time-crime type fixed effects. All standard errors are clustered by prosecutor-defendant.
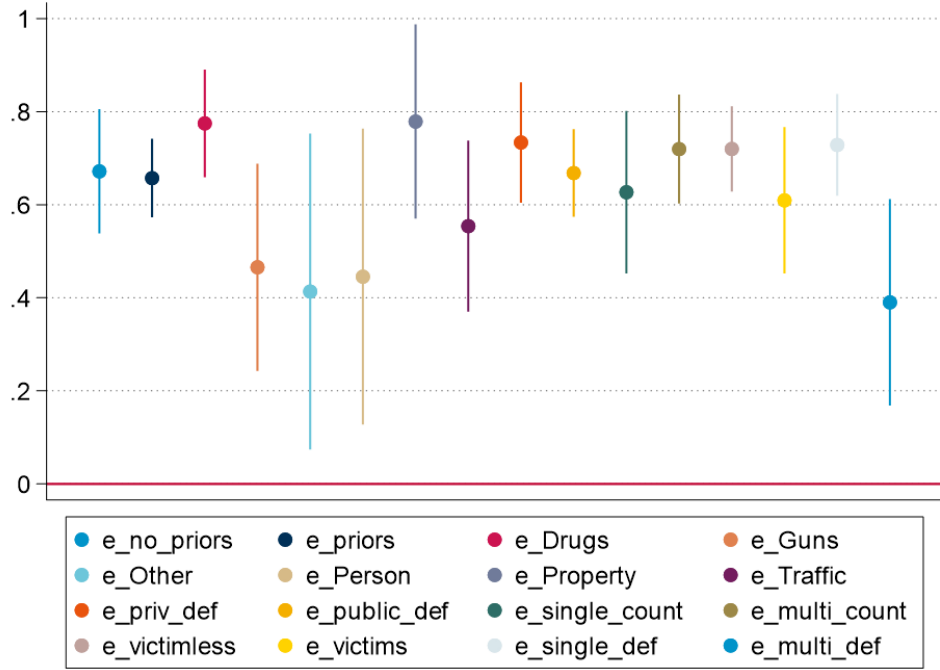
Figure 14: Plot of first-stage coefficients (and 95% confidence intervals) over 15 subsamples of data representing defendants with no prior convictions, more than one prior conviction, different types of charges (e.g. Drugs, Guns, Property, Traffic, Other), public defender, private defender, charges with no victims, charges with victims, single defendant charges, multi-defendant charges, single count charges, and multi-count charges.

D Comparing OLS with IV

Results from Figures 5-6 illustrate the considerable gap between the estimated effect using OLS and 2SLS (IV). This gap can be explained by an unobserved selection bias, measurement error, and / or heterogeneous treatment effects in the population. This section provides an empirical investigation as to what effect—if any—the latter is contributing to the observed differences in OLS and IV results. I first draw on the existing marginal-treatment-effect (MTE) literature that notes the LATE can be constructed by a weighted average of MTEs.[69] Specifically, one can recover the weights used to construct the LATE. The LATE framework includes four mutually exclusive groups of individuals:

---

[69]See Cornelissen et al. (2016) for an applied version of the technical work in (Heckman and Vytlacil, 1999, 2001, 2005; Heckman et al., 2006; Heckman and Vytlacil, 2007) among others.

1. **Never-takers:** individuals who would never be treated regardless of the instrument (e.g., defendants who would be prosecuted even if assigned to the most lenient prosecutor);

2. **Always-takers:** individuals who would always be treated regardless of the instrument (e.g., defendants who would not be prosecuted even if assigned to the least lenient prosecutor);

3. **Compliers:** individuals who are treated due to the instrument (e.g., a defendant who is not prosecuted (prosecuted) but would not have been not prosecuted (prosecuted) if assigned to a less (more) lenient prosecutor); and

4. **Defiers:** individuals who are treated in spite of the instrument (e.g., a defendant who is not prosecuted (prosecuted) but would not have been prosecuted (prosecuted) if assigned to a more (less) lenient prosecutor).

The monotonicity assumption rules out the existence of group (4); therefore, I assume the sample contains three types of individuals (never-takers, always-takers, and compliers). The LATE is generated exclusively from the treatment effect experienced by the compliers, whereas the OLS identifies the average treatment effect of the entire sample. Therefore, the difference in OLS and 2SLS could be driven by compliers being notably different across some observable dimensions from the rest of the population—and these differences drive heterogeneous treatment effects. To check if this effect heterogeneity across observables is driving the differences I observed, I re-run the OLS models but use the LATE weights derived in Appendix B3 in Heckman et al. (2006), Appendix C in Cornelissen et al. (2016), and Appendix A.6 in Andresen (2018). Operationally, I use the Stata package *mtefe* from (Andresen, 2018) to recover the weights. Essentially, these weights are higher for individuals who are more strongly affected by the instrument and thus more likely to be compliers. The tables on the following pages present comparisons of OLS and 2SLS estimates, as well as a

third column that represents weighted OLS estimates using the LATE weights. If heterogeneous treatment effects were the main driver of the larger IV estimates, I would expect the weighted OLS estimates to mirror the IV estimates. In fact, although the weighted OLS estimates are marginally larger in absolute value than the unweighted OLS estimates, they do not come close to the magnitude of the IV estimates. Therefore—as is the case in Agan et al. (2021)—prosecutors are on average not prosecuting defendants who pose a higher risk of subsequent arrests/convictions than the marginal defendants identified by the IV framework.

Future Re-arrest

OLS Re-weighted with LATE Weights: Black

VARIABLES	(1) 3 Months IV	(2) OLS	(3) W-OLS	(4) 6 Months IV	(5) OLS	(6) W-OLS	(7) 9 Months IV	(8) OLS	(9) W-OLS	(10) 12 Months IV	(11) OLS	(12) W-OLS	(13) 18 Months IV	(14) OLS	(15) W-OLS
Dismissal	-0.066	0.008	-0.002	-0.215	-0.008	-0.024**	-0.370*	-0.004	-0.031***	-0.335	-0.005	-0.044***	-0.484*	-0.023**	-0.068***
	(0.097)	(0.006)	(0.008)	(0.176)	(0.008)	(0.010)	(0.218)	(0.010)	(0.011)	(0.212)	(0.010)	(0.011)	(0.268)	(0.011)	(0.013)
Observations	11,233	11,233	11,214	11,233	11,233	11,214	11,233	11,233	11,214	11,233	11,233	11,214	11,233	11,233	11,214

Clustered standard errors at defendant-prosecutor level in parentheses

*p<.10; **p<.05; ***p<.01

Note. The dependent variable is a binary variable that takes the value of 1 if a defendant is arrested N months after the closure of his charge. All regressions include controls for the age of defendant, the type of crime, number of prior convictions, number of charges in the case the charge belongs to, if the charge is the top charge in the case, the number of victims, the number of co-defendants, the type of defense counsel, as well as team by year by month by weekend by broad crime-type fixed effects.

Future Re-arrest

OLS Re-weighted with LATE Weights: White

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	3 Months			6 Months			9 Months			12 Months			18 Months		
VARIABLES	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS
Dismissal	0.043	0.005	0.008	-0.052	-0.010	-0.009	-0.080	-0.016	-0.019	-0.111	-0.020	-0.023	-0.135*	-0.040**	-0.046***
	(0.058)	(0.013)	(0.014)	(0.066)	(0.016)	(0.017)	(0.074)	(0.017)	(0.019)	(0.078)	(0.017)	(0.019)	(0.080)	(0.015)	(0.017)
Observations	7,513	7,513	7,496	7,513	7,513	7,496	7,513	7,513	7,496	7,513	7,513	7,496	7,513	7,513	7,496

Clustered standard errors at defendant-prosecutor level in parentheses

*p<.10; **p<.05; ***p<.01

Note. The dependent variable is a binary variable that takes the value of 1 if a defendant is arrested N

months after the closure of his charge. All regressions include controls for the age of defendant, the type of

crime, number of prior convictions, number of charges in the case the charge belongs to, if the charge is the

top charge in the case, the number of victims, the number of co-defendants, the type of defense counsel, as

well as team by year by month by weekend by broad crime-type fixed effects.

Future Re-conviction
OLS Re-weighted with LATE Weights: Black

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	3 Months			6 Months			9 Months			12 Months			18 Months		
VARIABLES	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS
Dismissal	-0.096	0.007	0.001	-0.128	-0.003	-0.015*	-0.269	0.001	-0.020**	-0.279	-0.001	-0.032***	-0.452*	-0.010	-0.043***
	(0.094)	(0.005)	(0.008)	(0.124)	(0.007)	(0.009)	(0.170)	(0.008)	(0.010)	(0.168)	(0.007)	(0.010)	(0.234)	(0.009)	(0.011)
Observations	11,233	11,233	11,214	11,233	11,233	11,214	11,233	11,233	11,214	11,233	11,233	11,214	11,233	11,233	11,214

Clustered standard errors at defendant-prosecutor level in parentheses

*p<.10; **p<.05; ***p<.01

Note. The dependent variable is a binary variable that takes the value of 1 if a defendant is arrested N
months after their charge is closed and the arrest results in a conviction. All regressions include controls for
the age of defendant, the type of crime, number of prior convictions, number of charges in the case the charge
belongs to, if the charge is the top charge in the case, the number of victims, the number of co-defendants,
the type of defense counsel, as well as team by year by month by weekend by broad crime-type fixed effects.

Future Re-conviction

OLS Re-weighted with LATE Weights: White

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	3 Months			6 Months			9 Months			12 Months			18 Months		
	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS	IV	OLS	W-OLS
Dismissal	0.015	0.001	0.004	-0.057	-0.015	-0.014	-0.078	-0.019	-0.023	-0.086	-0.019	-0.024	-0.199**	-0.037***	-0.042***
	(0.051)	(0.010)	(0.011)	(0.057)	(0.013)	(0.014)	(0.068)	(0.014)	(0.015)	(0.079)	(0.013)	(0.016)	(0.080)	(0.010)	(0.012)
Observations	7,513	7,513	7,496	7,513	7,513	7,496	7,513	7,513	7,496	7,513	7,513	7,496	7,513	7,513	7,496

Clustered standard errors at defendant-prosecutor level in parentheses

*p<.10; **p<.05; ***p<.01

Note. The dependent variable is a binary variable that takes the value of 1 if a defendant is arrested N months after their charge is closed and the arrest results in a conviction. All regressions include controls for the age of defendant, the type of crime, number of prior convictions, number of charges in the case the charge belongs to, if the charge is the top charge in the case, the number of victims, the number of co-defendants, the type of defense counsel, as well as team by year by month by weekend by broad crime-type fixed effects.

E   Compliers Characteristics

Given the assumptions discussed in Appendix C, the examiner-design framework provides estimates of the causal effect of non-prosecution for compliers in my data.[70]   By definition, this effect is the difference in means for treated (non-prosecuted), Y_1, and untreated (prosecuted), Y_0 for the population whose treatment status D is affected by the instrument Z:

$$\mathbb{E}[Y_1 - Y_0 | D_1 - D_0 > 0] \ , \tag{16}$$

where $D_1 - D_0$ represents the change in treatment status after increasing the value of the instrument. Individuals for whom this difference is greater (equal) to zero are compliers (never or always takers). As derived in Imbens and Angrist (1994), 2SLS estimates the LATE using the following equation (using binary Z without loss for simplicity):

$$\mathbb{E}[Y_1 - Y_0 | D_1 - D_0 > 0] = \frac{\mathbb{E}[Y|Z = 1] - \mathbb{E}[Y|Z = 0]}{\mathbb{E}[D|Z = 1] - \mathbb{E}[D|Z = 0]} \ .$$

Similarly, Abadie (2002) uses the LATE assumptions to derive the average treated and untreated means for compliers:

$$\mathbb{E}[Y_1 | D_1 - D_0 > 0] = \frac{\mathbb{E}[YD|Z = 1] - \mathbb{E}[YD|Z = 0]}{\mathbb{E}[D|Z = 1] - \mathbb{E}[D|Z = 0]} \tag{17}$$

$$\mathbb{E}[Y_0 | D_1 - D_0 > 0] = \frac{\mathbb{E}[Y(1 - D)|Z = 1] - \mathbb{E}[Y(1 - D)|Z = 0]}{\mathbb{E}[(1 - D)|Z = 1] - \mathbb{E}[(1 - D)|Z = 0]} \ .$$

(18)

---

[70]See Appendix D for a full definition of compliers.

A quick way to empirically estimate equation 17 is to replace the dependent variable in the traditional 2SLS procedure (Y) with $W = Y \cdot D$. Similarly, equation E can be estimated by replacing the dependent variable in the traditional 2SLS procedure (Y) with $W = Y \cdot (1-D)$, as well as the treatment variable D with $1-D$. As shown in equation 16, 2SLS using either W identifies $\mathbb{E}[W_1 - W_0 | D_1 - D_0 > 0]$, which after substituting in each W, gives $\mathbb{E}[Y_1 | D_1 - D_0 > 0]$ and $\mathbb{E}[Y_0 | D_1 - D_0 > 0]$, respectively.[71] I then use these objects to estimate parameters in my simple model. Figures 15 and 16 illustrate the treated (non-prosecuted) and untreated (prosecuted) means within the complier subsample, compared with the sample-wide mean of all defendants (of the same race). For each look-ahead period, the LATE is simply the difference in the dismissed and prosecuted point estimate. For the following discussion, I discuss 18-month windows unless otherwise stated for clarity. First, when looking at the complier means for re-arrest probabilities (Figure 15) and re-conviction probabilities (Figure 16), salient racial differences stands out. Non-prosecuted and prosecuted White compliers are similar to non-compliers, with prosecution inducing a small insignificant gap increase in re-arrest / re-conviction probabilities. However, Black compliers exhibit starkly different patterns. To start, Black prosecuted compliers are *more* likely to be arrested and / or convicted than the prosecuted Black sample average. Additionally, Black non-prosecuted compliers are *less* likely to be arrested / convicted than the non-prosecuted Black sample average. These two trends suggest marginal Black defendants are particularly responsive to prosecution compared with the infra-marginal defendants. Also note no significant racial difference exists in the re-arrest / re-conviction probabilities for non-prosecuted compliers; therefore, the effect heterogeneity is driven purely by a substantial increase in re-arrest / re-conviction probability for prosecuted Black compliers.

---

[71] Hat tip to Peter Hull for revealing this trick from Angrist et al. (2013, 2022) in Lecture 3 of a workshop on IVs for Scott Cunningham's Mixtape Sessions.
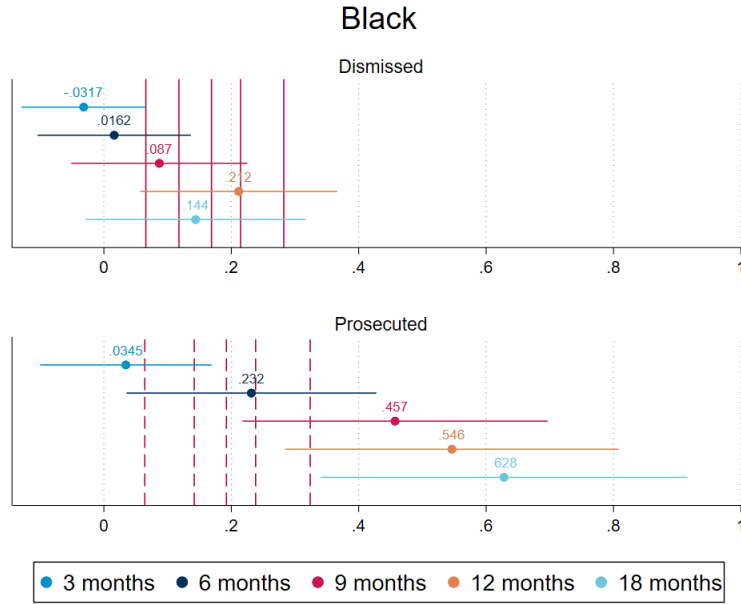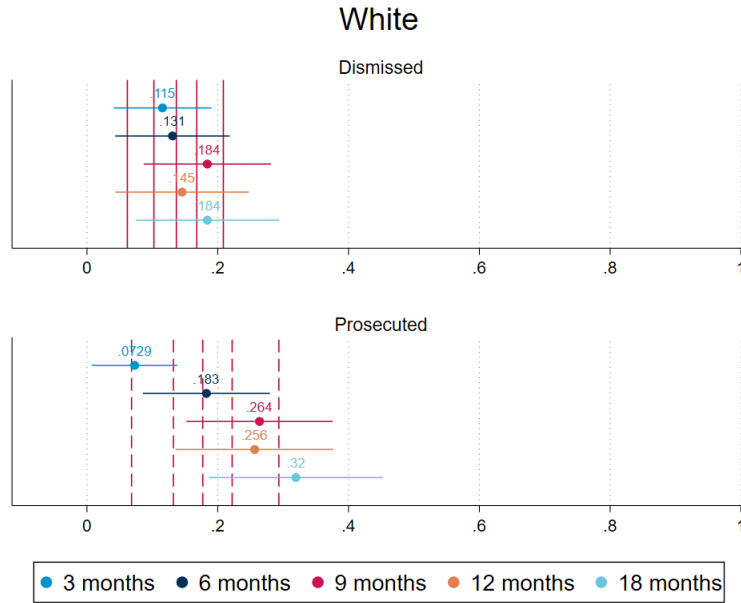
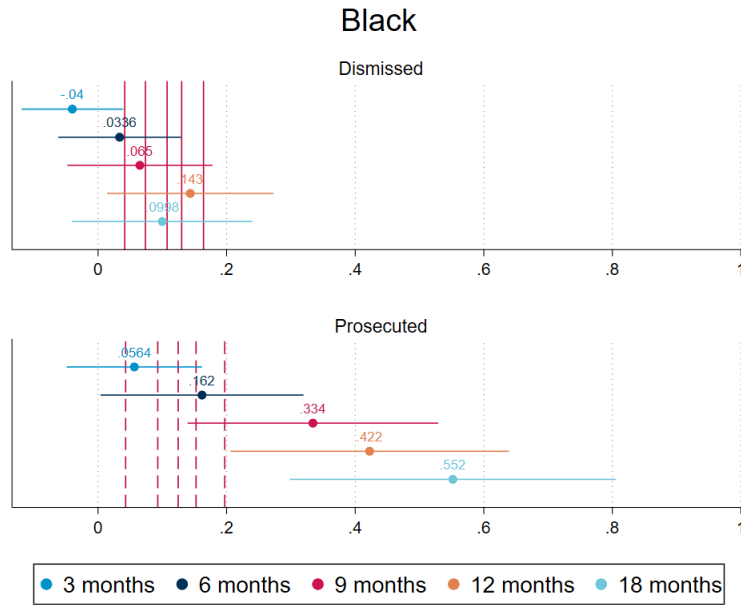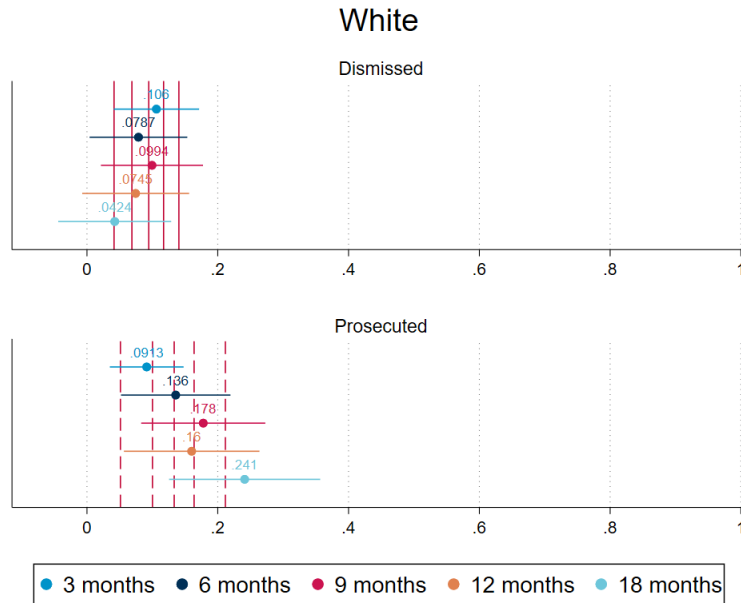Figure 15: Average re-arrest probabilities for (a) Black and (b) White non-prosecuted and prosecuted compliers. The solid (dashed) red lines—from left to right—represent the 3-, 6-, 9-, 12-, and 18-month re-arrest rates for all non-prosecuted (prosecuted) defendants. All regressions include the same controls in Figures 5 and 6.

Figure 16: Average probabilities of re-arrests that result in convictions for (a) Black and (b) White non-prosecuted and prosecuted compliers. The solid (dashed) red lines—from left to right—represent the 3-, 6-, 9-, 12-, and 18-month re-arrest rates that result in conviction rates for all non-prosecuted (prosecuted) defendants. All regressions include the same controls in Figures 5 and 6.

F   Rap Sheets

When making decisions, prosecutors rely heavily on a person's prior criminal behavior. My main dataset contains the number of prior convictions a defendant has. However, one could imagine the type of conviction matters when deciding a case. For example, if the current case pertains to drugs, an important statistic would be how many prior *drug* convictions the person has. In discussions with prosecutors from the jurisdiction, details such as this one are obtained by reviewing an arrestee's rap sheet.[72] One of the novel features of the data is that I was given access to the rap sheets available for prosecutorial review. Unfortunately, the rap sheets do not come in a machine-readable format. Specifically, they come in portable document format (.pdf). I received two types of pdfs from the jurisdiction: (1) "true" pdfs that contain all of the typed text in a structured format and (2) scanned images of the "true" pdfs. The former can be fed directly into the parsing tool, whereas the latter requires an additional processing step. This additional step consists of utilizing an optical character recognition (OCR) tool that converts images to text.[73] After all pdfs are converted into machine-readable text, I extract the necessary information using a parsing tool developed in python. This parsing tool uses an array of regular expressions to extract—with surprising accuracy[74]—the following fields for every criminal event that occurred in the state of interest:

For arrests:

- Date of arrest

- Short charge description

- Warrant number

For convictions:

---

[72]Rap sheets are constructed by a state agency that compiles individuals' criminal history by requesting information from the FBI as well as the state database.

[73]Specifically, I use a proprietary OCR tool provided by Azure: Azure Cognitive Services Computer Vision Read API.

[74]Cite accuracy from truth sample analysis

- Date of disposition

- Short charge description

- Short description of the disposition (e.g. conviction or non-conviction with sentence)

- Warrant number

After extracting this information, I hand match the short text descriptions of the arrest and conviction charges to the corresponding criminal code. The final output allows me to see all prior arrests and convictions for a defendant, labeled with a code that identifies the exact crime.

G   Pilot Screening Program

Starting on May 23, 2021, the office launched a pilot program to evaluate the feasibility and potential impact a screening unit could have on case outcomes. The head prosecutor hired a retired member of their staff to pre-screen a small subset of cases generated by a specific arresting agency. The types of cases that were eligible were determined by historical dismissal rates. This selection criterion was chosen because these case types were assumed to be the most likely "to have eventually been dismissed anyway", and therefore should be dealt with earlier to clear the docket for more serious cases. Operationally, as of September 28, 2022, the screening attorney had screened 568 charges, representing approximately 40% of all charges referred to the office of the specific arresting agency during that time.

Because this pilot was started before the creation of a standard operating procedure, the screening attorney has only been allowed to outright dismiss a case for a narrow set of reasons. After conversations with the screening attorney, I recognized that the screening attorney felt constrained by following the historic office practices to decide whether to dismiss a case. As such, I devised a survey to be used for each screened case to capture additional

information about the case, as well as the screening attorney's beliefs about prosecuting the case. Specifically, the survey asks the screening attorney for their underlying preferences for prosecution, by allowing three responses to the survey question: "What was your final decision on this case?":

1. I dismissed case;

2. I didn't dismiss, but the office should consider dismissing this kind of case in the future; and

3. I approved to prosecute.

To assess the potential impact of screening cases, I coded both dismissed charges (1) and those the screening attorney believes should be dismissed in the future (2) as dismissals and compared them with prosecuted charges (3), and then found the average by condition (screened v. unscreened) and race. To protect against the results being driven by unobserved case characteristics, I run the following simple model via OLS:

$$y_i = \beta SCREEN_i \cdot BLACK_i + \gamma BLACK_i + \alpha SCREEN_i + \boldsymbol{\theta X}_i + \epsilon_i \,, \qquad (19)$$

where y is the binary described above that obtains a value of 1 if the charge is dismissed or recommended for dismissal, $SCREEN$ indicates whether the charge, i, is screened (1) or unscreened (0), and $BLACK$ indicates whether the person accused of the charge is Black (1) or White (0). $\boldsymbol{\theta X}$ represents a vector of charge-level controls, including indicators for the number of prior convictions, number of concurrent charges, the age of the accused at the time of arrest, as well as the severity of the charge (measured by the average exposure of a defendant found guilty of the charge within my entire sample). The estimates of interest are β and α, which represent the impact of screening on the prosecution rates of Black and White males arrested by the specific agency. Because the implementation of the screening

program was unannounced to both residents and the arresting agency, we have no reason to believe the unobserved characteristics of the charges would be correlated with the start of the screening program; the estimates should be properly identified without the need for additional charge-level controls. Figure 17 graphically presents my estimates. The difference between the slopes of the light- and navy-blue lines represents the estimates for α and β, respectively. Specifically, I recover estimates of $\hat{\beta} = 8\%$ (robust standard error=3%, p-value=0.015) and $\hat{\alpha} = -3\%$ (robust standard error=4%, p-value=0.497). A salient result that can be recovered from both panels in Figure 17 is that screening has a heterogeneous impact on non-prosecution rates by race. Black males see an increase in non-prosecution from 16% to nearly 21% with the introduction of screening. By contrast, White men facing similar charges see no statistically significant change in their 13% dismissal rate (the point estimate actually drops to 10%). The similarity in the estimates of the raw (Figure 17 panel (a)) and conditional (Figure 17 panel (b)) models gives me further confidence that the timing of the screening program generated plausibly exogenous variation.

To further confirm these results, I also take advantage of the fact that charges from other arresting agencies were not eligible to be screened. First, I use these other agencies as control groups and use a difference-in-differences (DiD) framework where I include the same controls as above and recover similar estimates of $\hat{\beta}^{DiD} = 7\%$ (robust standard error=4%, p-value=0.103) $\hat{\alpha}^{DiD} = -5\%$ (robust standard error=3%, p-value=0.118). Finally, I use a synthetic control method to choose weights on the other agencies to best predict a counterfactual for the agency of interest and then compute the DiD estimate.[75] Here again, I obtain similar estimates of $\hat{\beta}^{SCM} = 9\%$ (robust standard error=3%, p-value=0.013) $\hat{\alpha}^{SCM} = -13\%$ (robust standard error=5%, p-value=0.008). In summary, I am confident of the results that imply implementing a screening process increased the non-prosecution rate for Black individuals and had no detectable (or slightly negative) effect on the non-prosecution rate for

---

[75]Here, I only use the race-specific dismissal rates to choose weights. For both Black and White samples, the algorithm chooses sparse weights (i.e., places zero weight on at least two of the six potential donor agencies.

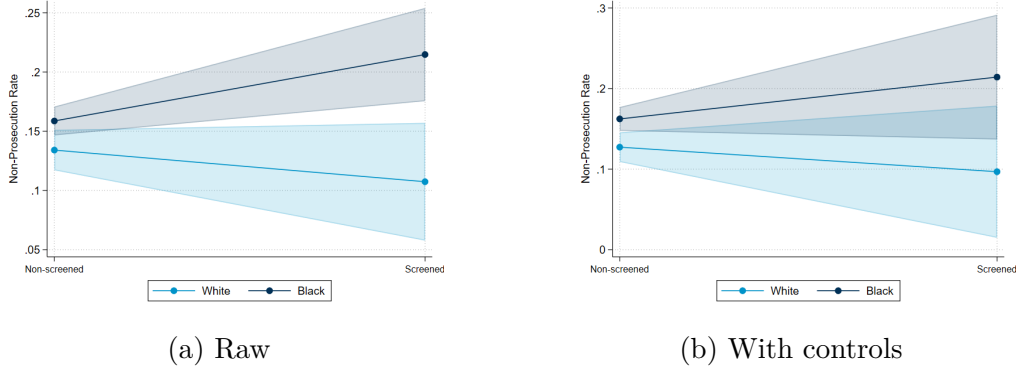|	(a) Raw	(b) With controls

Figure 17: Difference by race in non-prosecution rates for charges eligible for screening before (non-screened) versus after (screened) the screening process was implemented. Controls include indicators for the number of prior convictions, number of concurrent charges, age of accused at arrest, the severity of the charge (as measured by the average exposure of a defendant found guilty of the charge within my entire sample), and fixed effects for year and month of arrest. Sample includes only Black and White male arrestees arrested for screenable offenses from a single police department. N=5,576. Robust standard errors used to create shaded 95% confidence intervals.

White individuals.

H   Synthetic Population

In section 7, I utilized a synthetic population created by Research Triangle Institute (RTI) International (Wheaton, 2014). This synthetic population uses publicly available census microdata and assigns each record to a census block while respecting the marginal distributions of various aggregated census counts by census block group using a method called iterative proportional fitting. The variables used to match individual households to census blocks are age of the head of household, household income, household size, and race of heads of households. This synthetic population is then used as an input to estimate the predicted number of drug users and dealers at the block group level. However, my models for both drug use and sales require more information than is used to place households into block groups (e.g., marital status, personal income, employment status, education, and gender). RTI provides the individual identifier for each census record; therefore, these variables can be merged into the synthetic population dataset. A concern would be that the information

used to place individuals into census blocks does not have any predictive accuracy on the other characteristics being merged in. To ameliorate this concern, I aggregate the synthetic population to the block-group level for all dimensions being used in my model and compare it with the block-group aggregates provided by census. Figures 18 to 25 present these comparisons. As one can see, for every dimension, the synthetic population correlates highly with actual block-group rates provided by the Census.
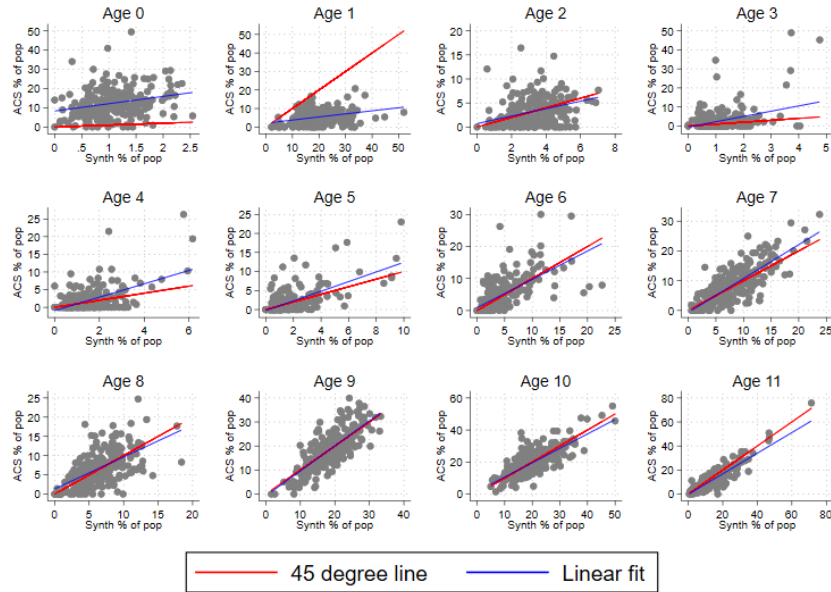

Figure 18: Comparison of aggregate synthetic population with ACS block-group aggregates by youngest (top left) to oldest (bottom right) age groups. Red line represents 100% accuracy; blue line represents population-weighted OLS fit line.

I   Modeling Drug Use and Sales

I begin by collecting data on self-reported use and sale of illicit drugs described in detail in section 4.2 from the National Drug Use and Health Survey (NSDUH). I first filter to include only responses from those living in small metro areas with fewer than 1 million inhabitants. After filtering the raw survey data, I expand by the supplied survey weights and draw a 10% random sample. I then create a test dataset by setting aside a random sample of 33% of
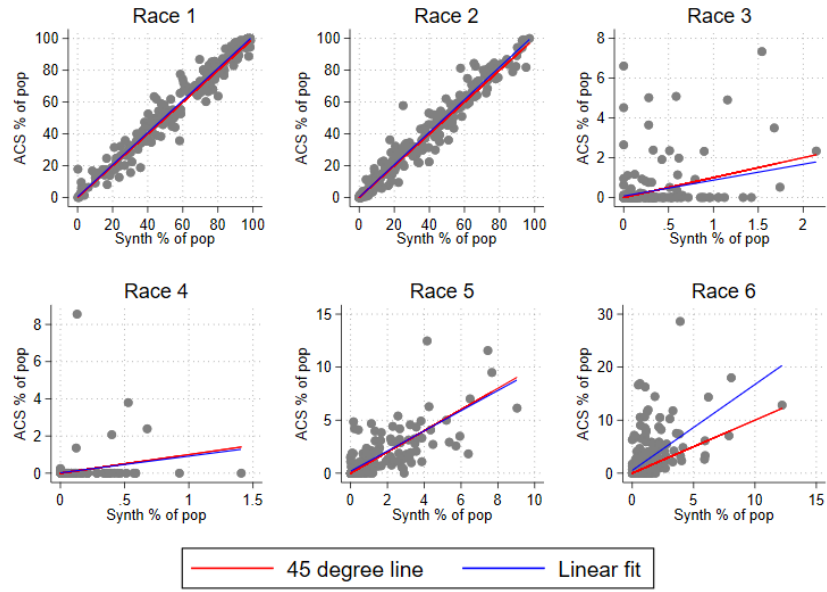
Figure 19: Comparison of aggregate synthetic population with ACS block-group aggregates by race. The categories of interest for this article are White (top-left panel) and Black (2nd from the left in the top row). Red line represents 100% accuracy; blue line represents population weighted OLS fit line.
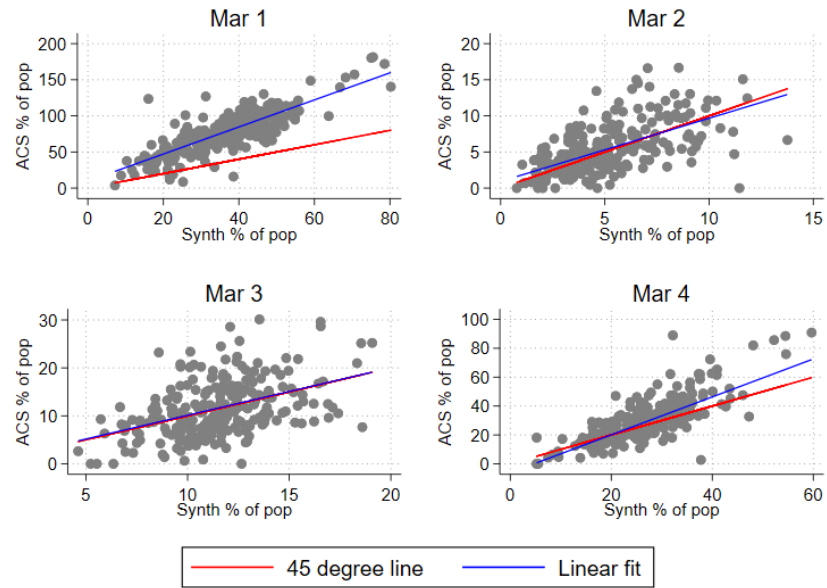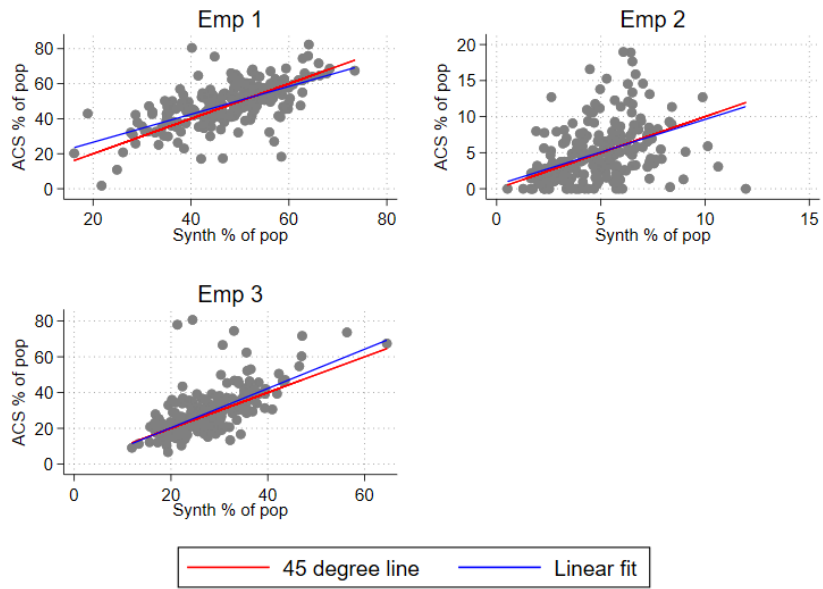

Figure 20: Comparison of aggregate synthetic population with ACS block-group aggregates by marital status. Red line represents 100% accuracy; blue line represents population-weighted OLS fit line.
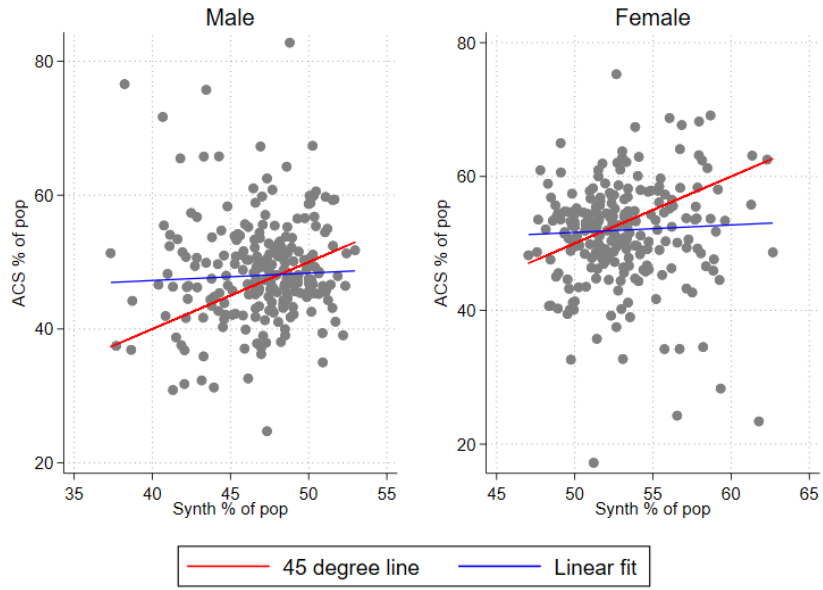
Figure 21: Comparison of aggregate synthetic population with ACS block-group aggregates by male (left panel) and female (right panel). Red line represents 100% accuracy; blue line represents population-weighted OLS fit line.

Figure 22: Comparison of aggregate synthetic population with ACS block-group aggregates by education. Red line represents 100% accuracy; blue line represents population-weighted OLS fit line.
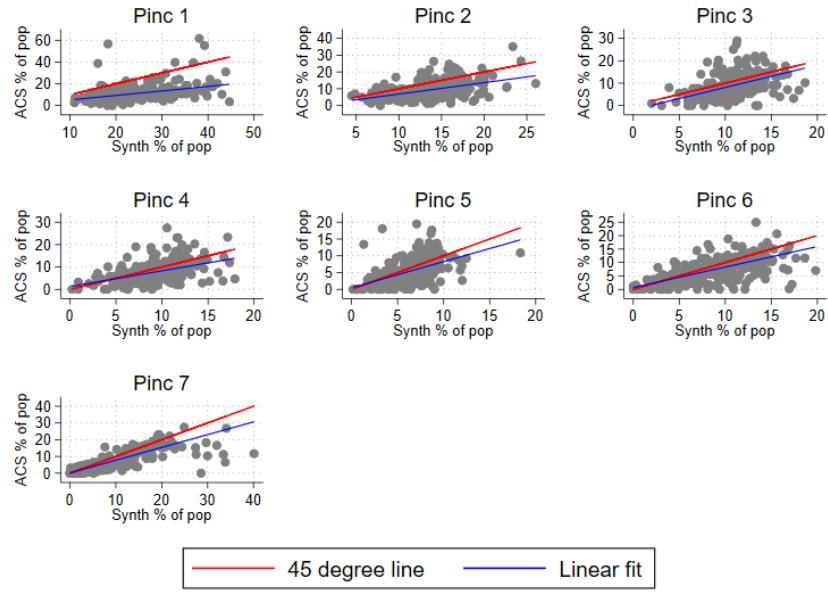
86

Figure 23: Comparison of aggregate synthetic population with ACS block-group aggregates by lowest personal income group (top-left panel) to highest (bottom-right panel). Red line represents 100% accuracy; blue line represents population-weighted OLS fit line.
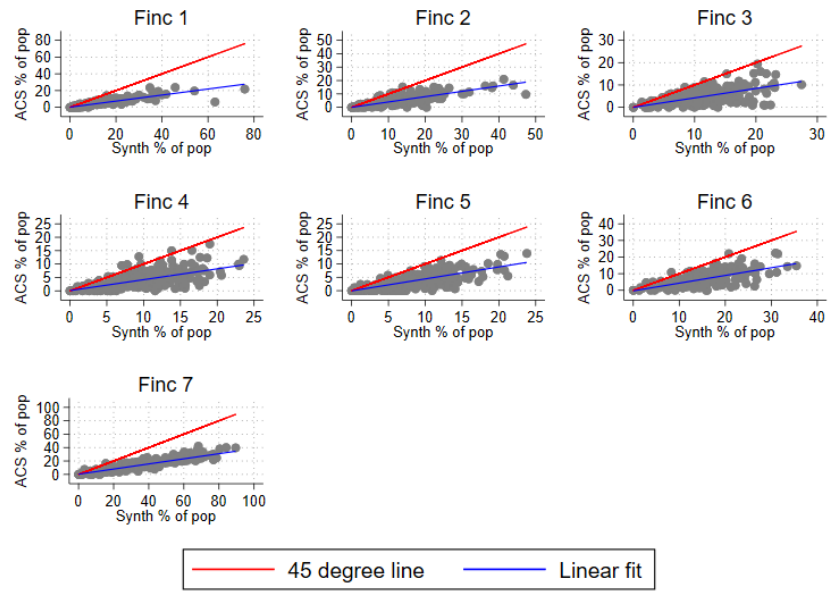

Figure 24: Comparison of aggregate synthetic population with ACS block-group aggregates by lowest family income group ($10K top left panel) to highest ($75K bottom right panel). Red line represents 100% accuracy, blue line represents population-weighted OLS fit line.
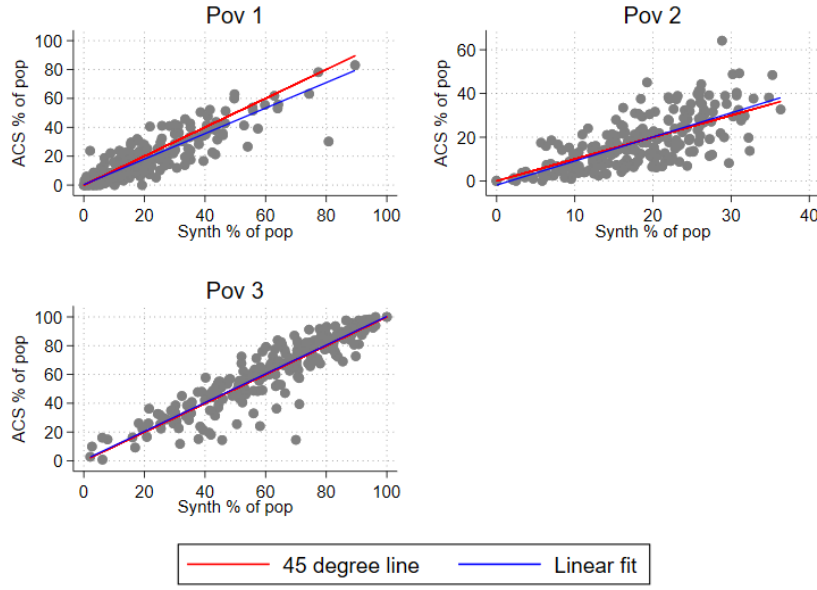
Figure 25: Comparison of aggregate synthetic population with ACS block-group aggregates by poverty status (in poverty is top right). Red line represents 100% accuracy; blue line represents population-weighted OLS fit line.

the full sample. I use this sample to test the accuracy and precision of my final prediction model.[76]

Drug use and drug sales are low-probability events resulting in an imbalanced dataset. This imbalance poses serious classification problems. Luckily, this issue is well documented and has well-thought-out solutions. I utilize a procedure described in Batista et al. (2003), which first creates synthetic observations that over-sample the minority class (drug users/sellers) using the synthetic minority over-sampling technique (SMOTE) then removes Tomek links from this synthesized sample. Although a full description of this technique is well beyond the scope of this paper, a short explanation is in order. To begin, SMOTE randomly selects an observation from the minority class and finds its five nearest neighbors using the features contained in the dataset (e.g., age, race, gender, income, etc.). Synthetic observations are then created using a convex combination (weighted average) of these data points. This pro-

---

[76]The following description of my modeling strategy is only applied to the 67% of the full sample not contained in this test dataset (training dataset).

cess is repeated until balance between minority and majority classes is achieved (i.e., the number of observations for drugs users is equal to the number of observations of non-drug users).[77] The next step is to identify *Tomek* links within the majority class and remove them.[78] Let $d(a, b)$ represent the Euclidean distance between unique observations a and b in the feature space, where both are from different classes. A *Tomek* link (a, b) exists if an observation $i \notin \{a, b\}$ that has a smaller distance between a or b than $d(a, b)$ does not exist. In other words, if a is b's closest neighbor *and* b is a's closest neighbor *and* both observations are of different classes, (a, b) is a *Tomek* link. These observations represent the "border-line" cases that are harder to classify, and therefore, removing them helps algorithms build stronger rules with less noise. The resulting sample consists of roughly equal observations of both classes, allowing more accurate prediction algorithms to be trained.

Table 6: Best Performing Models

Drug Sales				Non-marijuana Drug Use			
Raw							
Model	**Year**	**Stack Level**	**Number of Models**	**Model**	**Year**	**Stack Level**	**Number of Models**
NeuralNetFastAI	2015	2	53	NeuralNetFastAI	2015	2	95
NeuralNetFastAI	2016	2	60	NeuralNetFastAI	2016	2	95
NeuralNetFastAI	2017	2	53	NeuralNetFastAI	2017	2	95
NeuralNetFastAI	2018	2	53	NeuralNetFastAI	2018	2	95
NeuralNetFastAI	2019	1	14	NeuralNetFastAI	2019	2	53
Resampled							
RandomForestGini	2015	2	47	LightGBMLarge	2015	1	7
RandomForestGini	2016	2	47	XGBoost	2016	3	99
RandomForestGini	2017	2	47	RandomForestGini	2017	3	93
RandomForestGini	2018	2	47	LightGBMXT	2018	3	99
LightGBMXT	2019	2	123	XGBoost	2019	1	7

Note: Best model selected as model with highest $f1$ score in test data. Output from running *leaderboard()* function from AutoGluon over test data, with documentation here.

After resampling the data, I then turn to choosing the correct model to maximize the validity of the predictions. I turn to the off-the-shelf tool AutoGluon to allow the data to pick the model best suited for my task. I fed the training data into this function and allowed it to run for 12 hours for each year (2015-2019) and each task (non-marijuana drug use, marijuana drug use, and drug sales) using either raw training data or re-sampled training data.[79] Because no *a priori* asymmetry is present in the cost of making either type I (i.e.,

---

[77]See Chawla et al. (2002).
[78]See Tomek (1976).
[79]The models were trained using a desktop PC with AMD Threadripper 3960X 3.8 GHz 24-Core Processor

predicting a user/seller when they are not) or type II error (i.e., predicting non-use/seller when they in fact are using/selling), I utilize the F1 metric to train all models. This metric is standard in the literature and puts equal weight on the false-negative and false-positive rates. The F1 metric is calculated by finding the harmonic mean of precision, p, and recall, r:

$$F_1 = \frac{2}{r^{-1} + p^{-1}}, \tag{20}$$

where recall is the number of true/correct positives identified by the model divided by the number of all positives in the sample and precision is the number of true/correct positive positives identified by the model divided by the number of all positives identified by the model—including those not identified correctly. Table 6 summarizes the models selected for each year. Of note, surprising homogeneity exists in model selection year over year.

Throughout the paper, I utilize models of drug use and sales to answer questions regarding race. Therefore, understanding the performance of the model by race is important. For example, in the most extreme case, if the false-positive rate differed considerably by race, any inference made about the racial disparity in drug use/sales would be thrown into question. Figures 26 and 27 present confusion matrices by race using raw and re-sampled data. Comparing the sensitivity—how well the model correctly predicts positives (in this case drug use/sales)—with specificity—how well the model correctly predicts negatives (i.e., non-drug use/sales)—is instructive. To start, the sensitivity of the model when predicting drug use for Black and White individuals using raw (re-sampled) training data is 98.9% (96.0%) and 98.4% (87.8%), respectively, whereas specificity of the model when predicting drug use for Black and White individuals using raw (re-sampled) training data is 86.2% (99.8%) and 65.8% (95.9%), respectively. Similarly, for drug sales, the sensitivity of the model for Black and White individuals using raw (re-sampled) training data is 99.4% (99.0%) and 99.8% (97.0%), respectively, whereas specificity of the model when predicting drug sales for Black
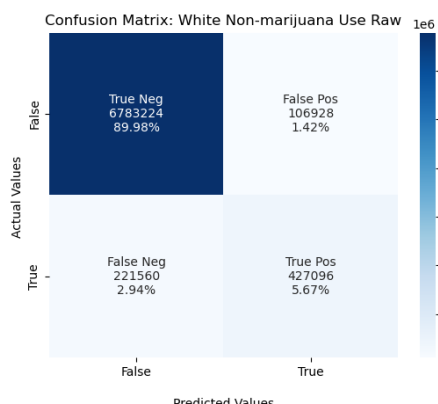

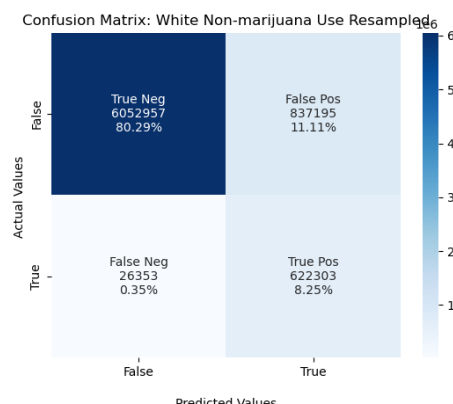
with 128 GB DDR4 RAM and a NVIDIA 3080 Ti GPU.

90
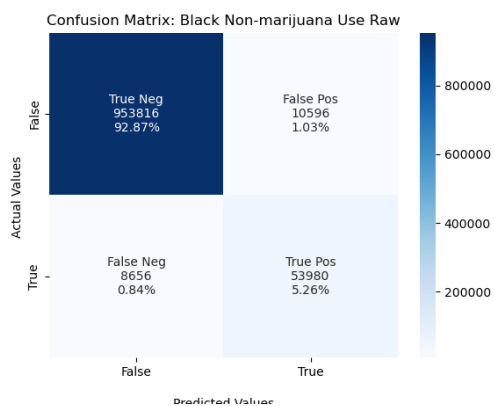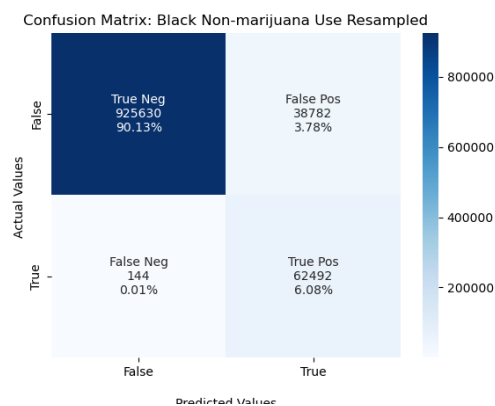
Figure 26: Confusion matrices for all drug-use predictions made by models aggregated over 2015-2019 for (a)/(c) Black and White individuals using models trained on raw training data and (b)/(d) Black and White individuals using models trained on resampled training data
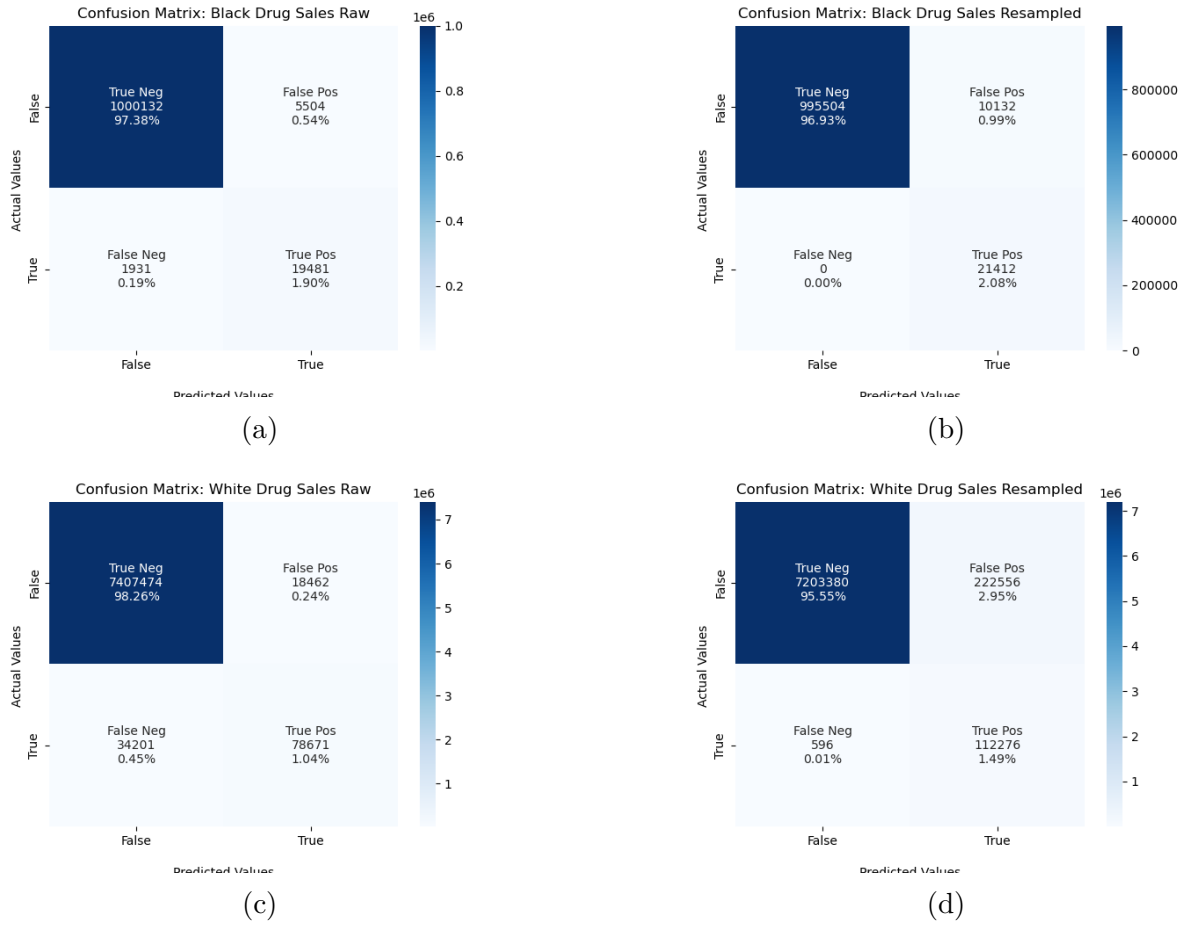
Figure 27: Confusion matrices for all drug-sales predictions made by models aggregated over 2015-2019 for: (a)/(c) Black and White individuals using models trained on raw training data and (b)/(d) Black and White individuals using models trained on resampled training data

and White individuals using raw (re-sampled) training data is 91.0% (100%) and 69.7% (99.5%), respectively. A salient trend is that resampling greatly increases the ability of the model to correctly identify drug users and sellers, with the trade-off being more false positives. Another interesting pattern is that in all cases, the model performs better for the Black population, with this finding being less pronounced in the models trained on re-sampled models. I see these trends as interesting in their own right and look forward to following this line of inquiry in future papers.

J   SafeGraph

When studying the racial disparities found in arrests per capita, one may worry that rather than biased policing, racial differences may exist in the nature of drug use. A common story is that Black individuals are more likely to use and/or sell drugs in outdoor areas or areas with high visibility, and thus are more likely to be detected by police patrols. To build a proxy for the "publicness" of an area, I construct relative foot-traffic estimates by census-block group. This foot-traffic data were graciously provided by SafeGraph.[80] To be specific, I collected 463,183 monthly observations from January 2018 to March 2020 and from June 2021 to August 2022 from 17,815 places whose location fell within the jurisdiction of interest.[81] This approach left me with 114 million visits to places during this time. The median distance the median visitor traveled from their residence to the place they were visiting was approximately six miles, and the median of the median dwell times was 34 minutes. Therefore, these data represent short trips to stores and other places of interest. To obtain an exact measure, I calculated the average normalized visits by census-block group as a proxy for the amount of foot traffic in a block group relative to other block groups in the

---

[80]SafeGraph is a data company that aggregates anonymized location data from numerous applications to provide insights about physical places, via the SafeGraph Community. To enhance privacy, SafeGraph excludes information from census-block groups if fewer than five devices visited an establishment in a month from a given census-block group.

[81]I dropped all months during which a statewide emergency was declared due to COVID-19.

area.[82] My assumption is that block groups with more visitors are more likely to represent public places, where drug use and sales would be easier to detect. Further work could be done to test alternative proxies using these data; I leave this research to future projects.

---

[82]Visits were normalized by the total number of visits in the state for the month. See SafeGraph's monthly patterns documentation for more details.