NBER WORKING PAPER SERIES

RACIAL DIVISIONS AND CRIMINAL JUSTICE: EVIDENCE FROM SOUTHERN STATE COURTS

Benjamin Feigenberg
Conrad Miller

Working Paper 24726
http://www.nber.org/papers/w24726

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2018, Revised September 2019

We thank Amanda Agan, Ernesto Dal Bo, Fred Finan, Anil Jain, Jonathan Leonard, Abhishek Nagaraj, Aurelie Ouss, Evan Rose, Yotam Shem-Tov, and seminar participants at Dartmouth College, Pomona College, UC Berkeley, UCLA, University of Illinois at Chicago, University of Chicago Crime Lab, Chicago Harris, Duke, and Brown for comments. We thank the Center for Science and Law for providing access to Alabama court data, and we thank Claire Lim, James Snyder, Jr., and David Strömberg for generously sharing voting data they collected on state ballot propositions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Racial Divisions and Criminal Justice: Evidence from Southern State Courts
Benjamin Feigenberg and Conrad Miller
NBER Working Paper No. 24726
June 2018, Revised September 2019
JEL No. J15,K14,K41,K42

ABSTRACT

The US criminal justice system is exceptionally punitive. We test whether racial heterogeneity is one cause, exploiting cross-jurisdiction variation in punishment in four Southern states. We estimate the causal effect of jurisdiction on arrest charge outcome, validating our estimates using a quasi-experimental research design based on defendants charged in multiple jurisdictions. Consistent with a model of in-group bias in electorate preferences, the relationship between local punishment severity and black population share follows an inverted U-shape. Within states, defendants are 27%-54% more likely to be sentenced to incarceration in 'peak' heterogeneous jurisdictions than in homogeneous jurisdictions.

Benjamin Feigenberg
Department of Economics
University of Illinois at Chicago
601 South Morgan Street
706 University Hall
Chicago, IL 60607
bfeigenb@uic.edu

Conrad Miller
Haas School of Business
University of California, Berkeley
Berkeley, CA 94720
and NBER
ccmiller@berkeley.edu

1 Introduction

The United States incarcerates residents at a higher rate than any other country in the world. While less than five percent of the world's population resides in the US, nearly 25% of the world's prison population is held in US facilities (Walmsley, 2016). Though differences in violent crime rates can in part explain this pattern, the US is also exceptionally punitive (Pfaff, 2014). Some observers have argued that race plays a key role in driving American criminal justice policy (Alexander, 2010). There is prima facie evidence: US blacks are incarcerated at six times the rate of whites and face longer sentences for similar crimes (Carson, 2014; Rehavi and Starr, 2014). Race may play a broader role, even influencing the incarceration rate for US whites, which itself would rank near the top among developed nations (Gottschalk, 2015). Just as racial heterogeneity predicts lower support for redistribution and public goods (Alesina et al., 1999), it may increase support for harsher punishment if, for example, voters prefer to punish outgroup members more severely. In this paper, we ask whether racial heterogeneity can in part explain US exceptionalism in criminal justice.

Empirical research on the role of race in criminal justice policy is complicated by the difficulty of separating the relative importance of policy versus underlying criminal conduct in generating cross-country variation in incarceration rates. Harmonized micro data covering the US and a significant number of other countries do not exist, and differences in the definitions of crimes across countries would make harmonization difficult. Instead, we study the relationship between racial divisions and criminal justice policy by investigating cross-jurisdiction variation in punishment *within* US states. In doing so, we take advantage of harmonized data and fixed criminal codes within states and exploit the substantial within-state variation in how criminal law is enforced.

While much statutory criminal justice policy is driven by state-level legislation, localities have significant discretion in how they enforce those laws, and that discretion is tied to electorate preferences. Prosecutors and judges are often locally elected and influence outcomes at each stage of the criminal justice process: prosecutors decide what charges to file; prosecutors negotiate plea bargains; judges make sentencing decisions after conviction. The electorate may also affect adjudication outcomes by serving as jurors or influencing spending on indigent defense. A 2016 New York Times Upshot article illustrates the role of local politics in driving local punishment severity with a quote from the elected prosecutor in Dearborn County, Indiana: "I am proud of the fact that we send more people to jail than other counties...My constituents are the people who decide whether I keep doing my job. The governor can't make me. The legislature can't make me" (Keller and Pearce, 2016).

In this paper, we evaluate the role that racial heterogeneity plays in determining criminal justice outcomes. We first estimate local *punishment severity*, the causal effect of jurisdiction on the outcome of a criminal arrest charge, using data from four Southern states. We then link variation in punishment severity to local racial heterogeneity in the population. Consistent with a simple model of ingroup bias in electorate preferences, we find that the relationship between local punishment severity and black population share follows an inverted U-shape: jurisdictions with the largest white and black shares are relatively lenient while intermediate, heterogeneous jurisdictions are more punitive.

To measure punishment severity, we use rich criminal justice administrative data that track criminal charges from arrest through sentencing, including dropped charges. Our benchmark measures of punishment severity are the jurisdiction fixed effects we estimate in a regression of charge outcomes on a rich set of covariates, including defendant demographics and criminal history, the specific arrest charge, and the year of the charge. To validate our benchmark estimates, we develop and apply a 'mover'-based quasi-experimental research design that exploits variation in outcomes for defendants arrested in multiple jurisdictions (Chetty et al., 2014). We show that our benchmark punishment severity estimates accurately predict the *within-defendant* changes in charge outcomes coinciding with changes in jurisdiction. Throughout the analysis, our benchmark specifications focus on the share of charges that lead to incarceration sentences (the *confinement rate*) as the relevant measure of punitiveness, though we present supplementary analyses that confirm that our findings are similar if we use conviction or sentence length as the outcome or employ case-level rather than charge-level specifications.

The data cover charges from 2000 to 2014 in Alabama, North Carolina, Texas, and Virginia, with ranges varying by state. These states account for about 20% of all prisoners held under state jurisdiction in the US. We focus on the South because there is substantial variation in racial composition across Southern counties. Across counties in the states we study, the black share of the population ranges from 0% to 85%. In all four states, district attorneys are locally elected; in all but Virginia, judges are locally elected. The data reveal substantial within-state heterogeneity in jail and prison admissions: across states, the coefficient of variation for jurisdiction-level admissions per capita ranges from 38% to 72% in our sample.

The variation we measure in admissions per capita is matched by substantial heterogeneity in punishment severity: the coefficient of variation for punishment severity ranges from 25% to 57%. A defendant charged in a jurisdiction in the top quartile by punishment severity is 1.8 to 3.6 times more likely to be incarcerated for a given charge than the same defendant charged in a jurisdiction in the bottom quartile. We find that 80%-93% of the difference in confinement rates between top and bottom quartile jurisdictions is explained by the causal effect of jurisdiction. Interestingly, punishment severity estimates constructed separately by race are highly correlated. Jurisdictions that are more punitive for black defendants are also more punitive for white defendants.

We next document the relationship between local punishment severity and racial heterogeneity. We motivate our analysis with a simple model of ingroup bias where voters prefer more severe punishment when offenders are more likely to belong to a different racial group. Prior work documents that common group membership is associated with declines in envy and punishment for misbehavior as well as increases in charitable concerns and rewards for good behavior (Chen and Li, 2009). This mechanism implies that the relationship between local punishment severity and the black share of the population (or share of defendants) will follow an inverted U-shape; while white voters prefer more punitive policy as the black share of defendants increases, for jurisdictions with sufficiently large black populations, the pivotal voter is more likely to be black and to support less harsh punishment.

Lacking a natural experiment that generates variation in racial composition across jurisdictions, we test for an inverted U-shape pattern in the cross-section and adjust for other jurisdiction covariates. The predicted relationship is borne out in the data and the magnitude of the relationship is large. Our regression results imply that punishment severity peaks where the black share of the population (defendants) is around 0.3 (0.4). At this peak, predicted confinement rates for a given offense are 24 (43) log points larger than predicted confinement rates for the same offense in a jurisdiction with a black share of the population (defendants) that is zero. If we adjust for other jurisdiction characteristics, notably population density, the difference between 'peak' heterogeneous and homogeneous jurisdictions is reduced but remains substantial at 14-27 log points. By contrast, *within* jurisdictions, the 'unexplained' black-white gap in confinement rates varies from 11-19 log points across states.[1] We find a similar inverted U-shape relationship between contemporaneous punishment severity and a jurisdiction's slave share of the population in 1860. Notably, we do not find evidence of non-monotonic relationships between punishment severity and other jurisdiction characteristics, and we estimate that selection on unobservables would have to be substantially larger in magnitude than selection on observables to explain the cross-sectional relationship between punishment severity and black share.

Surprisingly, lagged growth in violent crime, which has been previously identified as an important driver of cross-state variation in incarceration rates (Western, 2006), is uncorrelated with punishment severity in our sample, and contemporary crime rates are *negatively* correlated with severity. Following the cross-state analysis presented in Western (2006), we introduce additional covariates in alternative specifications, including average income, income inequality, the fraction of prime-aged males in the population, and population density. Among these covariates, population density consistently predicts higher confinement rates. Consistent with Cohen and Yang (2019), we also find that Republican vote share predicts harsher punishment. To provide support for the hypothesis that local racial composition affects punishment severity through the preferences of the local electorate, we use data on local voting for statewide ballot measures aimed at increasing pun-

[1] Rehavi and Starr (2014) find a 10% unexplained gap in sentence length in federal courts.

ishment harshness or limiting the rights of the accused (Lim et al., 2015). We find that electoral support for these measures predicts more severe punishment and also has an inverted U-shaped relationship with the local black share.

We conclude by simulating outcomes under a counterfactual in which all jurisdictions adopt the punishment severity imposed by those jurisdictions with racially homogeneous populations. Under this counterfactual, we show that overall confinement rates and racial confinement rate gaps fall by approximately 15%, on average, once we account for both the static effect of lower punishment severity on confinement outcomes and the dynamic effect of lower punishment severity on defendants' criminal histories.

Our work contributes to a political economy literature that studies the association between local racial composition and policy preferences. Alesina et al. (1999) provide evidence that public goods spending is inversely related to ethnic fragmentation in US cities and argue that this finding is driven by cross-group policy preference heterogeneity. Luttmer (2001) shows that self-reported support for welfare spending is increasing in the share of local recipients from the respondent's own racial group. The link between ethnic fragmentation and support for redistribution is also buttressed by Dahlberg et al. (2012), which shows that plausibly exogenous increases in immigration to Swedish municipalities are associated with decreases in support for redistribution. We argue that the inverted U-shape relationship between black population share and severity of incarceration policy in our data can be explained by the same racial ingroup bias that drives the positive association between racial homogeneity and support for redistribution.

In emphasizing the importance of racial divisions as a key driver of electoral preferences and local punitiveness, we build on a large literature that highlights the racialized nature of crime policy in the US (Muhammad, 2010) and the role of 'racial threat' in explaining policy and punishment preferences (Key, 1949; Glaser, 1994; Enos, 2015; Unnever and Cullen, 2007). The racial threat literature studies how the presence of racial and ethnic minority populations affects white voting behavior and policy preferences. While findings in this literature are generally inconsistent, the most recent and compelling evidence suggests that a larger minority population increases white voter turnout and support for conservative policies and candidates (Enos, 2015).

A related literature uses survey data to link white racial attitudes to criminal justice preferences. Whites who express more racial resentment are more likely to support capital punishment and other harsh crime-control policies (Unnever and Cullen, 2010). Respondents who are primed to consider the prison population as 'more black' express more concern about crime and greater acceptance of punitive policies (Hetey and Eberhardt, 2014). There is evidence that public support for 'tough on crime' policies tracks national incarceration rates over time (Enns, 2014). While we cannot measure local preferences directly, we measure local policy in the form of punishment severity.[2]

[2] As noted above, we also analyze data on local voting for criminal justice-related statewide ballot measures, a

Motivated by the racial threat hypothesis, several papers test for a relationship between state racial composition and imprisonment rates, with mixed results. Most relevant to our work, Keen and Jacobs (2009) finds an inverted U-shaped relationship between black population share and racial *disparities* in state prison admissions per capita. In contrast with this past research, we focus on county-level criminal justice and develop a mover-based identification strategy to credibly isolate the causal effect of charge location on sentencing. We find an inverted U-shaped relationship between county black population share and punishment severity that applies to *all* defendants, including both black and white defendants.

Our findings provide a potential explanation for a pattern that has been documented in several recent papers: courts and police officers appear to be more punitive in areas with larger black and Hispanic populations. Rehavi and Starr (2014) find that, conditional on other case and defendant characteristics, sentences are more severe in federal districts where black defendants are concentrated. Raphael and Rozo (2017) find that California police officers are more likely to formally book (rather than cite or release without punishment) arrested juveniles in cities with relatively large black and Hispanic populations. Goncalves and Mello (2018) find that Florida police officers are more punitive in writing speeding tickets in locations where black and Hispanic drivers are concentrated. Each of these papers is focused on measuring racial disparities in outcomes, and finds that observed gaps decrease with the inclusion of locality fixed effects. By contrast, we are focused on estimating unbiased measures of locality punishment severity itself, and our research design is suited for this objective. Moreover, while the papers above provide evidence that local punitiveness is positively correlated with the black share of the local population, a key prediction of our ingroup bias model is that the relationship is *non-monotonic*: local punishment severity is increasing in the share of the black population over some range, but then decreasing as the black population share continues to rise. We document this non-monotonic relationship empirically.

A key feature of our data is that they include arrest charges that are dropped by prosecutors and convictions that do not result in incarceration sentences. By contrast, many past studies of racial disparities in sentencing use data that only include convictions that lead to incarceration sentences.[3] This restriction introduces selection bias: charges that result in conviction and incarceration sentences may differ in unobservable ways from charges that do not, and this selection may vary by jurisdiction. Indeed, we find that variation in conviction rates is an important source of variation in punishment severity. Moreover, we show that we would reach dramatically different conclusions about how punishment severity varies across jurisdictions if our data were limited to

noisy proxy for local punishment preferences.

[3] See, for instance, Miethe (1987). It is important to note that the extent of selection bias may be more limited in federal criminal cases than in state cases. Fischman and Schanzenbach (2012), for instance, conditions on conviction but argues that associated selection bias is limited because acquittals account for only one percent of the federal criminal cases that they analyze.

charges that lead to an incarceration sentence. On this point, a closely related paper is Rehavi and Starr (2014), which uses data tracking federal criminal cases from arrest through sentencing. The authors find that, conditional on arrest charge, a prosecutor's initial court charge is an important driver of racial disparities in sentencing. In particular, prosecutors are more likely to file charges carrying mandatory minimum sentences for black defendants.[4]

Our paper builds on a literature that studies the role of electoral pressure on the composition and behavior of judges and, to a lesser extent, prosecutors. Broadly, this literature finds that judge and prosecutor behavior is tied to local electorate preferences. Voters elect like-minded public officials, or public officials are responsive to electorate preferences. Huber and Gordon (2004) and Berdejó and Yuchtman (2013) find that judges in Pennsylvania and Washington sentence serious crimes more severely when they come up for reelection. The authors argue that this pattern of judge behavior is driven by re-election incentives and the preferences of the electorate. Lim (2013) finds that judges in counties using partisan judicial elections exhibit different sentencing patterns from judges in counties using referendum judicial elections, and attributes these differential patterns to differences in electoral pressure under the two systems. Lim et al. (2015) finds that newspaper coverage increases sentence length by nonpartisan elected judges for violent crime, and argues this relationship is mediated through electorate preferences. There is also evidence that prosecutors respond to constituent preferences (Dyke, 2007; Nelson, 2014). In our model, the predicted relationship between local punishment severity and racial composition that we document is mediated through electorate preferences.

Our empirical strategy builds on past work that uses changes in workplace or residence to separate effects of individual characteristics from location-based factors. Our approach is most closely related to Chetty et al. (2014) who validate measures of teacher value-added using teacher moves from one school to another. We also adapt specification checks developed by Card et al. (2013, 2016) to evaluate worker-firm wage decomposition models (Abowd et al., 1999). Our paper is similar in spirit to Finkelstein et al. (2016) who decompose geographic variation in Medicare spending into location and patient effects by exploiting patient migration across markets, and then correlate estimated location and patient effects with observable characteristics.

The remainder of the paper is structured as follows. Section 2 describes the data used for the analysis. Section 3 discusses our approach to characterizing cross-jurisdiction differences in punishment severity, including our mover-based validation strategy, and provides estimates. Section 4 presents a model of racial ingroup bias to highlight the role that racial divisions may play in explaining this variation and empirically tests the predictions of the model. In Section 5 we simulate counterfactual confinement rates. Section 6 concludes.

[4] The issue of dropped charges is less relevant in the federal context. Over 90% of cases in Rehavi and Starr (2014) lead to conviction. By contrast, depending on the state, 22%-61% of charges are dropped in our data.

2 Data

We use administrative criminal justice data from four states: Alabama, North Carolina, Texas, and Virginia.[5] The data source and years of data we analyze for each state are presented in Table 1. We summarize the content of the data here and discuss data construction and state-specific institutional context in greater detail in Appendix A.

One key distinction across states is the data source. The data from Alabama, North Carolina, and Virginia are administrative court records, and include relatively detailed and complete information on criminal charges starting from the time they are filed in court. In principle, a limitation of these data is that they do not include information on criminal charges prior to court filing. Fortunately, in these states *all* arrests based on probable cause result in court charges, so we effectively have data on all valid arrests.[6] The Texas data are maintained by the Texas Department of Criminal History, and include data from arresting agencies (e.g. police departments), prosecutors, and courts. These data contain records for all qualifying arrests, including arrests that did not lead to a court charge. However, the data contain less detailed information on court processes, and do not identify whether a charge was ever filed in court.

Though the data from each state differ in their exact content, they all track state felony and misdemeanor criminal charges from arrest through sentencing, and share important data elements. Critically, data for all states include arrest charges that are ultimately dropped. Data from all states include information on each criminal charge, including the original arrest charge, the date of arrest, the court where the charge is assigned, final court charge, charge disposition, and, if the charge results in conviction, the final sentence. The data allow us to group charges into cases. Defendant information includes date of birth (except Virginia, which does not include year of birth), gender and race. Data from North Carolina and Texas also identify Hispanic defendants.

For all states, the data include property, violent, and drug offenses. We refer to offenses in these categories as 'core' offenses. The data also include 'crimes against society', including driving while intoxicated (DWI), writing bad checks, and trespassing. For all states, we drop non-DWI traffic offenses. We also exclude charges in which the final listed disposition is an intermediate

[5] We have also analyzed data from Arkansas and Maryland. However, we omit data from these states due to data quality issues. Including data from these states does not substantively affect any of the results reported in this paper.

[6] In these states, if an officer serves an arrest warrant or makes a warrantless arrest based on probable cause, the officer takes the arrested person before a magistrate. For a warrantless arrest, the magistrate determines whether there is probable cause for arrest. Once probable cause is determined, the magistrate sets conditions of release and issues the arrested person a court date for a first appearance before a court judge. At this stage, a court record is generated. Based on conversations with numerous court and law enforcement officials in each of these states, our understanding is that this process generally occurs without the involvement of prosecutors, except for some exceptionally high-profile cases. Note that, in some other states, after probable cause is determined prosecutors decide whether to file court charges.

outcome, such as a transfer between district and circuit courts or across jurisdictions. Lastly, we exclude technical probation and parole violations that do not result in new criminal charges. While we include all remaining charges in the baseline analysis, we also explore limiting charges to core offenses as a robustness check.

We drop charges for defendants aged below 16, which are likely to be adjudicated within the juvenile justice system. We also exclude offenses with fewer than 100 occurrences in the data. These offenses are rare—this restriction removes many specific offense codes from the data, but only around 1% of charges. Lastly, we drop offenses that by statute cannot lead to an incarceration sentence and offenses with zero instances that result in confinement.[7] This leaves us with about 400-600 unique offenses in each state.[8]

In Alabama and Virginia, we restrict to black and white defendants. In Alabama, American Indian-, Asian-, and Hispanic-coded defendants account for less than 0.25% of charges. In Virginia, the same categories amount for about 2% of charges. In North Carolina and Texas, we restrict to black, white, and Hispanic defendants.[9] American Indian- and Asian-coded defendants account for less than 2% and 1% of charges in these states, respectively. In all states, we drop defendants with missing race codes. These account for about 1% or less of charges in all states. See Appendix A for more details.

We generally define jurisdictions within states based on prosecutor and judge electoral districts. In all but North Carolina, the most granular partition among prosecutor and judge electoral districts is the county. For consistency, we use the county as our measure of jurisdiction for all states, but results are similar if we group counties into prosecutor or judge electoral districts.[10]

2.1 Confinement Sentence as a Benchmark Outcome Measure

There are several potential outcomes to use for measuring punishment severity. A criminal charge can be pursued or dropped by the prosecution. Pursued charges can result in conviction, acquittal, deferred judgment, or some other outcome. Conviction can lead to probation or confinement sentences of varying lengths, or an alternative sentence.

[7] Based on Sec. 12.23 of the Texas Penal Code, we thus exclude Class C misdemeanors in Texas, which represent the lowest level criminal offense and do not have jail or prison penalties. Similarly, based on 18.2-11 in the Code of Virginia, we exclude Class 3 and 4 misdemeanors in Virginia, which represent the lowest level criminal offenses and do not have jail or prison penalties.

[8] When we analyze court outcomes at the case level rather than the charge level as described below, we include excluded offenses when constructing controls if they are not the primary charge in the case.

[9] In Texas, Hispanic status is defined separately from race. We treat Hispanic as a distinct category, and re-define the black and white racial categories to only include non-Hispanic black and white defendants. In North Carolina, Hispanic is already defined as a distinct category.

[10] In North Carolina, 100 counties are divided into 43, 44, and 50 district court judge, prosecutor, and superior court judge districts.

Table 1: Data by State

State	Source	Year
Alabama	Alabama Administrative Office of the Courts	2000-2010
North Carolina	North Carolina Administrative Office of the Courts	2007-2014
Texas	Texas Department of Public Safety	2000-2010
Virginia	Virginia's Office of the Executive Secretary	2006-2014

Notes: Data sources are discussed in more detail in Appendix A.

For our measure of severity, we examine whether a given charge results in a jail or prison confinement sentence. This excludes alternative sentences, such as probation or suspended sentences, where the defendant may serve time in jail or prison if they violate the terms of their alternative sentence. We focus on confinement as our outcome given our particular interest in US exceptionalism in incarceration policy. We study the extensive margin of confinement rather than sentence length in part because our data generally do not include information on the mapping between nominal sentence length and realized sentence, which may vary across jurisdictions in ways we cannot measure.[11]

As a robustness check, we also examine two alternative outcomes, conviction and (nominal) sentence length. As we show below, results are qualitatively similar for all three outcomes.

2.2 Descriptive Statistics

We tabulate descriptive statistics for charge data from each state in Table 2. We include information on defendant demographics, charge characteristics, and charge outcomes. The number of charges in our data ranges from 1.9 million in Alabama to 5.9 million in Texas. The number of charges per defendant ranges from 2.3 in Texas to 3.1 in North Carolina. Across states, 71.1% to 78.7% of charges are filed against male defendants. Defendants are disproportionately black; while the black share of the population ranges from 11.8% in Texas to 26.1% in Alabama, the black share of defendants ranges from 24.4% in Texas to 43.1% in Virginia. In both Texas and North Carolina, the Hispanic share of defendants is *lower* than the Hispanic share of the population. However, there is evidence that law enforcement may underreport Hispanic status (Collister, 2015). Twenty-eight percent to 40% of charges are felonies. The distribution of offense types varies across states, though in each state a plurality of charges is for 'Other' offenses.

Note that charge outcomes vary significantly across states. In Texas, 40.2% of charges result in confinement, and 28.1% result in a jail or prison sentence of at least 90 days; in North Carolina, those shares are 8.4% and 3.6%. This is due in part to variation in severity across states, but may

[11] The same nominal sentence in two counties may lead to different realized sentences, for example, due to parole board decisions. Notably, parole board members are not locally elected.

Table 2: Charge-Level Descriptive Statistics

	Alabama	North Carolina	Texas	Virginia
Male	71.1	75.2	78.7	72.5
Black	37.2	42.3	24.4	43.1
Hispanic		4.2	30.5	
Age	32.9	31.6	31.0	
	(10.9)	(11.9)	(10.8)	
Felony	35.5	27.6	31.2	39.8
Property	17.0	31.0	22.1	33.6
Violent	10.0	13.6	12.2	11.1
Drug	17.3	19.5	21.8	15.0
Other	55.6	35.9	43.9	40.4
Dropped	40.4	61.1	22.3	43.4
Convicted	57.5	36.7	55.4	51.7
Probation	27.8	15.6	30.9	11.4
Confinement	21.2	8.4	40.2	18.9
Sentence $\geq$ 90 Days	16.1	3.6	28.1	9.5
N Defendants	727,419	1,840,251	2,588,641	1,108,911
N Charges	1,854,208	5,742,283	5,876,448	2,613,297
N Cases	1,221,317	3,984,894	4,931,314	1,777,549
Charges per Defendant	2.5	3.1	2.3	2.4
	(4.3)	(5.4)	(2.4)	(4.1)
Cases per Defendant	1.7	2.2	1.9	1.6
	(2.0)	(2.7)	(1.7)	(1.7)
Charges per Case	1.5	1.4	1.2	1.5
	(1.8)	(1.6)	(0.6)	(2.0)

Notes: Missing values reflect characteristics that are unavailable for particular states. 'Other' offenses include crimes against society and offenses we are unable to classify due to miscoding.

also be due to differences in charging behavior across states. Across states, the same crime may result in a different set of arrests, which may in turn result in a different set of recorded charges.[12] Throughout the analysis we focus on comparing jurisdictions within states.

We compare jail and prison admissions across jurisdictions within states in Table 3. We use three measures: jail and prison admissions per 100,000 residents (age 15 or above), jail and prison admissions per case, and the share of charges that lead to a jail or prison sentence. Throughout, we refer to the last measure as the *confinement rate*. While the first measure incorporates variation in number of cases and charges per capita across jurisdictions, the second and third measures come closer to capturing how a given case or charge is treated differently across jurisdictions.

There is substantial variation in all three measures. For admissions per 100,000 residents, the (unweighted) coefficient of variation varies from 38% in North Carolina to 72% in Texas. For admissions per case, the coefficient of variation varies from 26% in Virginia to 52% in Alabama. For confinement sentence per charge, the coefficient of variation varies from 29% in Virginia to 59% in Alabama.[13]

3 Estimating Punishment Severity

We posit that jurisdictions vary in their *punishment severity*—-they vary systematically in how they punish equivalent charges, so that there is a causal effect of jurisdiction on charge outcomes. A key objective of this paper is to measure and compare punishment severity across jurisdictions. Punishment severity reflects variation across jurisdictions in prosecutor and judge behavior, defense attorney quality, and jury preferences. The local electorate plays an important role by electing prosecutors and judges, serving as jurors, and by indirectly determining the level of funding for indigent defense.

To form our benchmark estimates of punishment severity, we estimate linear regression models where the dependent variable is the outcome of a charge and the explanatory variables are rich observable charge and defendant characteristics and jurisdiction fixed effects. Specifically, we estimate models of the following form, *separately by state*:

$$y_{ict} = \tau_{cth(i,t)} + x_i\gamma^x + z_{it}\gamma^z + \theta_{j(i,c,t)} + \epsilon_{ict} \tag{1}$$

where i indexes individuals, c indexes initial charge, t indexes year, $h(i,t)$ is the criminal history for individual i at time t, and $j(i,c,t)$ is the court jurisdiction. y_{ict} is an indicator for any con-

[12] Charging behavior may vary across jurisdictions within states, an issue we explore in more detail below.

[13] Variation in admissions per case and confinement sentence per change is not due to chance; if we randomly allocate cases to jurisdictions, maintaining the number of cases per jurisdiction, the coefficient of variation ranges from 2% to 6%.

Table 3: Jail and Prison Admissions Across Jurisdictions

	Alabama	North Carolina	Texas	Virginia
Admissions Per 100,000:				
Mean (Weighted)	604	599	787	726
Mean	626	569	511	766
SD	(419)	(219)	(366)	(481)
Admissions Per Case:				
Mean (Weighted)	0.251	0.116	0.406	0.236
Mean	0.223	0.107	0.232	0.234
SD	(0.118)	(0.033)	(0.119)	(0.062)
Confinement Sentence Per Charge:				
Mean (Weighted)	0.213	0.085	0.408	0.193
Mean	0.189	0.077	0.236	0.191
SD	(0.111)	(0.024)	(0.120)	(0.056)
N Jurisdictions	67	100	253	118

Notes: 'Admissions per 100,000' is the total number of cases resulting in a jail or prison sentence in a county and year divided by county population that is age 15 or above in that year, averaged across years, and multiplied by 100,000. 'Admissions per Case' is the rate that *cases* result in a jail or prison sentence. 'Confinement Sentence Per Charge' is the rate that *charges* result in a jail or prison sentence. Weighted means are weighted by jurisdiction population in 2000.

finement sentence, our primary charge outcome of interest; $\tau_{cth(i,t)}$ are specific arrest offense code by defendant criminal history by year fixed effects; x_i is a vector of time invariant individual controls (defendant race and gender); z_{it} is a vector of time-varying individual controls (age). Finally, $\theta_{j(i,c,t)}$ is a jurisdiction fixed effect, which we use to construct our punishment severity measure.

Our objective is to measure the causal effect of each jurisdiction on charge outcomes. Equation (1) includes rich controls; there are 400-600 unique arrest offense codes per state and several criminal history categories, which we describe below. However, there remain unobserved determinants of charge outcomes that we cannot measure, e.g. the quality of the evidence possessed by the prosecutor. For equation (1) to recover the causal effects of interest, it must satisfy a *selection on observables* assumption: conditional on $\tau_{cth(i,t)}$, x_i, and z_{it}, unobserved determinants of charge outcomes must be uncorrelated with jurisdiction. We will probe this assumption further in Section 3.2, where we use defendants arrested in multiple jurisdictions to validate our baseline punishment severity estimates. Further note that we model punishment severity as additively separable from other charge characteristics. That is, we assume that jurisdictions that are punitive for one type of charge (e.g., a violent crime) are also punitive for other types of charges (e.g., a property crime). We assess this assumption in Section 3.1.1.

To construct criminal history $h(i,t)$ we rely on state-specific sentencing legislation that defines mandatory or suggested sentencing enhancements based on the severity of current charges in combination with the number and severity of prior convictions.[14] The number of resultant categories ranges from two (for misdemeanor defendants in Texas) to 20 (for defendants charged with larceny in Virginia). A more detailed description of state-specific criminal history classification is provided in Appendix B. Though the criminal history classifications are in some instances quite coarse (particularly for misdemeanor defendants in Texas and North Carolina), we have verified that results are robust to defining criminal history based on federal statute.[15] This alternative approach generates a more continuous measure that incorporates number of past convictions, number of past incarceration sentences, and length of past incarceration sentences and allows for a consistent criminal history classification across states.

To construct punishment severity from the θ_j estimates, we add a state-specific constant so that the result is the predicted confinement rate for each jurisdiction using the overall composition of charges in that state.[16] This procedure ensures that log transformations of punishment severity

14 In Virginia, enhancements also depend on incarceration history.

15 This is a point system based on prior offenses. For each prior offense, a defendant receives: 3 points if the sentence was longer than 390 days, 2 points if the sentence was longer than 60 days, and 1 point for a conviction. Defendants are assigned to 1 of 6 categories depending on total prior points. The six categories consist of defendants with 0-1 points, 2-3 points, 4-6 points, 7-9 points, 10-12 points, and 13 or more points.

16 That is, we average predicted values for each charge in that state derived from (1) but omitting the jurisdiction effect *corresponding to the location of the charge*, and then add the estimated jurisdiction effect, θ_j, to construct the punishment severity for jurisdiction j.

are well-defined, which we use when making cross-state comparisons of punishment severity in Section 4.

The coefficient estimates for equation (1) are presented in Table 4. The pattern of coefficients is consistent with past research (for example, Rehavi and Starr, 2014). Conditional on offense charge, criminal history, year, and jurisdiction, black and male defendants are more likely to receive confinement sentences. Where the data are available, Hispanic defendants are also more likely to receive confinement sentences. The relationship between punishment and defendant age is non-monotonic, increasing in age at younger ages and decreasing at older ages.

Punishment severity estimates are summarized in Table 5 and displayed on state maps in Figure 1. Notably, controlling for observable offense and defendant characteristics does not substantially mute cross-jurisdiction variation in confinement rates.[17] Table 5 also includes the average punishment severity for jurisdictions in the top and bottom quartiles of jurisdictions, ranked by punishment severity. The differences in punishment severity between quartiles is substantial. Across states, defendants are 1.8 to 3.6 times more likely to face a confinement sentence in fourth quartile jurisdictions than in first quartile jurisdictions.

3.1 Robustness Checks and Extensions

In this section we assess the robustness of our benchmark punishment severity estimates in several ways. First, we scope the potential for 'match effects'— interactions between charge characteristics and punishment severity. Second, we assess whether variation in the mapping of crimes to arrests can account for the variation in punishment severity we observe. Third, we analyze arrests at the case level rather than the charge level. Fourth, we estimate punishment severity using alternative charge outcomes: conviction and sentence length.

3.1.1 Match Effects

Our estimating equation (1) models punishment severity as separable from other charge characteristics such as crime type or defendant race. This may obscure heterogeneity in punishment severity across types of charges or defendants. For example, a jurisdiction that we characterize as moderately punitive may be lenient with property crimes but harsh with violent crimes. We gauge whether such match effects are empirically important. To do this, we re-estimate punishment severity separately for different types of charges: by defendant race (black versus white), by

[17] Variation in estimated punishment severity is not due to chance; if we randomly allocate cases to jurisdictions, maintaining the number of cases per jurisdiction, the standard deviation of pseudo punishment severity ranges from 0.2% in NC to 1.7% in Texas. Outside of Texas, there are more than 600 charges in each jurisdiction. In Texas, there are 25 counties with fewer than 600 charges, and 10 with fewer than 100. Excluding these counties from the analysis has no meaningful effect on any of the results presented in this paper.

Table 4: Coefficient Estimates from Punishment Severity Models

Outcome: Confinement	Alabama	North Carolina	Texas	Virginia
Black	0.033**	0.020**	0.072**	0.027**
	(0.001)	(0.000)	(0.000)	(0.001)
Hispanic		0.033**	0.056**	
		(0.001)	(0.000)	
Male	0.045**	0.027**	0.100**	0.040**
	(0.001)	(0.000)	(0.000)	(0.001)
Age	0.001**	0.005**	0.013**	
	(0.000)	(0.000)	(0.000)	
Age^2 $\times$ 100	-0.002**	-0.005**	-0.015**	
	(0.000)	(0.000)	(0.000)	
Criminal History $\times$ Charge $\times$ Year Fixed Effects	✓	✓	✓	✓
Jurisdiction Fixed Effects	✓	✓	✓	✓
N Charges	1,854,208	5,742,283	5,876,448	2,613,297
Adjusted R^2	0.153	0.093	0.187	0.163
Mean Confinement	0.212	0.084	0.402	0.189

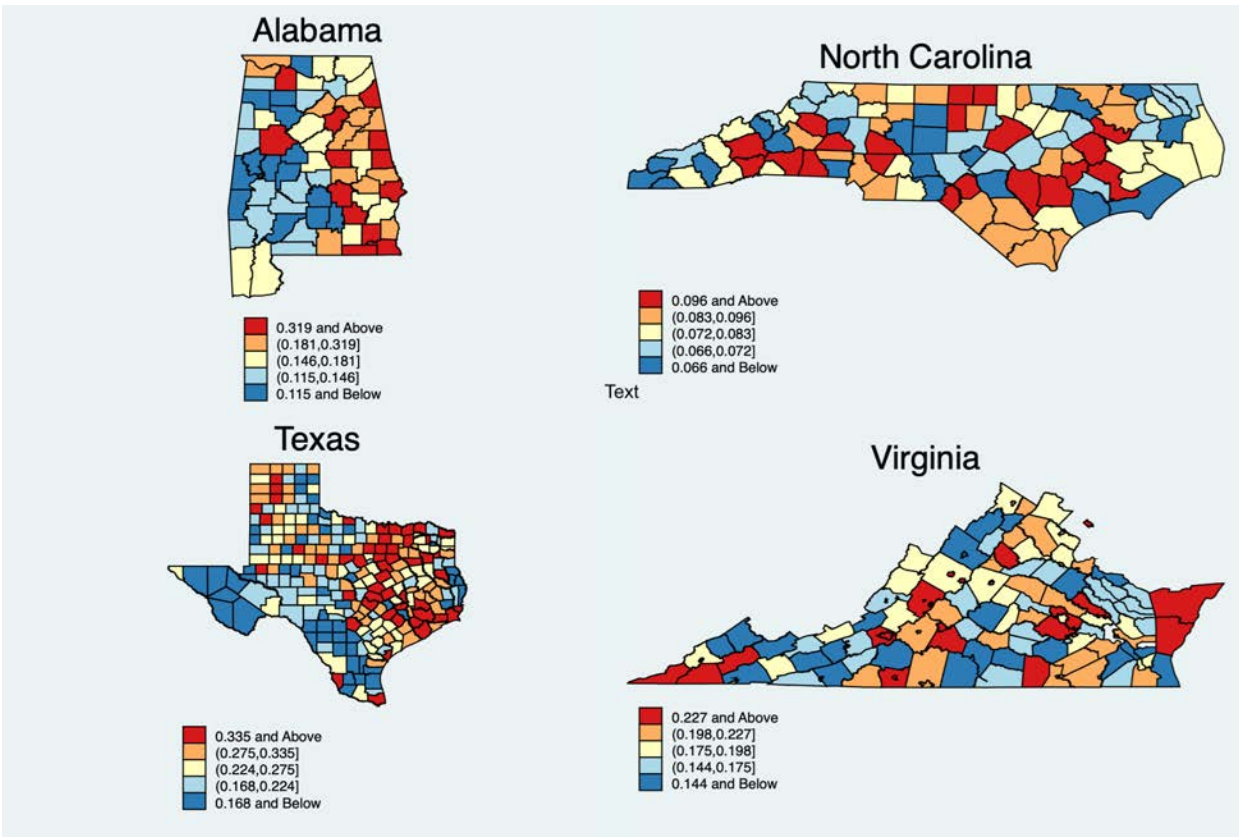
Notes: Table presents coefficients from state-specific estimates of equation (1). Missing values reflect characteristics that are unavailable for particular states.
Standard errors clustered by defendant in parentheses. ˜ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table 5: Summary of Baseline Punishment Severity Estimates

	Alabama	North Carolina	Texas	Virginia
Avg. Confinement Rate (%)	18.9	7.7	23.6	19.1
SD of Punishment Severity	10.7	2.0	11.2	4.8
Number of Jurisdictions	67	100	253	118
Adjusted Q1 Rate	10.3	6.1	13.2	13.3
Adjusted Q4 Rate	37.0	10.8	41.3	25.6

Notes: Punishment severity estimates are derived by estimating equation (1) separately by state and then adding a state-specific constant as described in Section 3. The outcome is an indicator for any confinement sentence. Further details on the estimation of punishment severity are discussed in Section 3.

Figure 1: Maps of Punishment Severity



Note: These maps depict punishment severity estimates by county. Punishment severity estimates are derived by estimating equation (1) separately by state and then adding a state-specific constant as described in Section 3. The outcome is an indicator for any confinement sentence. Further details on the estimation of punishment severity are discussed in Section 3.

criminal history (first-time versus repeat offenders), and by crime category. We estimate punishment severity separately for property, violent, and drug charges, and for those three core categories pooled together. We then compare estimates across subsamples.

Correlations are presented separately by state in Panel A of Table 6. The correlation between punishment severity for black and white defendants ranges from 0.78 to 0.95. The correlation between punishment severity for first-time and repeat offenders ranges from 0.83 to 0.92. Punishment severity does not vary significantly by defendant race or criminal history.

The correlation between punishment severity estimates based on all and core offenses ranges from 0.89 to 0.97. Punishment severity is similar whether or not we restrict to core offenses. The correlations between specific crime categories are generally smaller, ranging from 0.60 to 0.87. The one exception is the correlation between violent and drug crime-based estimates in Virginia, which is 0.36.

As an alternative approach to assessing match effects, we follow Card et al. (2013) and compare the adjusted R^2 of our baseline model equation (1) and a more saturated model that includes: (a) interactions between jurisdiction effects and crime type and (b) interactions between jurisdiction effects and an indicator for whether a charge is a defendant's first offense. Saturating the baseline model increases the adjusted R^2 by a modest 3%-16% across states, implying a limited role for match effects based on crime type or criminal history.

In summary, we find that jurisdictions that are punitive for one type of defendant or charge are also punitive for other types. Moreover, we will show that the patterns in punishment severity that we document below are quantitatively similar for each subcategory of charges.

3.1.2 Selection into Arrest and Arrest Charge

Another measurement concern that could bias cross-jurisdiction comparisons is that the threshold that determines whether (a) an arrest is made and (b) which specific charge is filed may vary across jurisdictions. For example, some police departments may be more lenient than others in deciding whether to arrest a suspect. In that case, jurisdictions with fewer marginal arrests may appear more severe in part because the composition of offenses that actually lead to an arrest may be (unobservably) more serious. Among arrests, some police departments may pursue more severe charges, conditional on the underlying criminal conduct. Because we control flexibly for the initial court charge as our measure of underlying conduct, jurisdictions with more (unobserved) charge upgrading by police officers may consequently appear less punitive in part because the composition of offenses that actually lead to a given initial charge may be (unobservably) less serious. We address selection into arrest and selection into specific arrest charge in turn.

To evaluate selection into arrest, we investigate how a proxy for selection into the court data correlates with estimated punishment severity. In Section 4.2, we also try to control for this selec-

Table 6: Subsample and Alternative Outcome Correlations

	Alabama	North Carolina	Texas	Virginia
Subsample Correlations:				
Black vs. White	0.95	0.80	0.78	0.82
First vs. Subsequent Offense	0.92	0.88	0.88	0.83
All vs. Core	0.89	0.97	0.96	0.97
Property vs. Violent	0.77	0.74	0.75	0.60
Property vs. Drug	0.83	0.79	0.82	0.65
Violent vs. Drug	0.87	0.67	0.77	0.36
Granular vs. Coarse	0.99	0.98	0.99	0.98
Alternative Outcome Correlations:				
Confinement vs. Conviction	0.64	0.52	0.53	0.51
Confinement vs. Sentence $\geq$ 90 Days	0.83	0.72	0.38	0.81
Confinement vs. Sentence Length	0.97	0.95	0.60	0.96
Confinement vs. Cond. Sentence Length	-0.03	-0.19	-0.23	-0.06
Number of Jurisdictions	67	100	253	118

Notes: In the top panel, jurisdiction-specific punishment severity is constructed separately for each referenced subsample of defendants or charges. There are two jurisdictions where we are unable to calculate punishment severity for black defendants due to insufficient data (King County, Texas and McMullen County, Texas). There is one jurisdiction where we are unable to calculate punishment severity for violent crimes due to insufficient data (King County, Texas). In the bottom panel, we correlate baseline punishment severity estimates with punishment severity estimates derived using the following alternative charge outcomes: an indicator for conviction; an indicator for a jail or prison sentence at least 90 days; the inverse hyperbolic sine of sentence length, recorded in days; the inverse hyperbolic sine of sentence length, recorded in days, restricted to charges that results in a confinement sentence.

tion when measuring the relationship between punishment severity and jurisdiction characteristics. To proxy for selection, we calculate the ratio of charges in the court data for a given county and year to crimes reported in the FBI Uniform Crime Reports (UCR) for the same county and year, and then average that ratio across years by county. We restrict to Part I crimes reported in the UCR data: arson, aggravated assault, burglary, murder, rape, robbery, and theft. This excludes drug and public order offenses, which make up a significant share of offenses. However, as documented in Section 3.1.1, our punishment severity measures are highly correlated across crime categories.[18]

Within states, the correlation between punishment severity and the charge to crime ratio is -0.20 in Alabama, -0.18 in North Carolina, -0.08 in Texas, and -0.21 in Virginia. Jurisdictions that we measure as more punitive also have somewhat fewer recorded charges relative to the number of reported crimes. Reassuringly, when we include the charge to crime ratio as a control variable below, it has little effect on the estimated relationship between punishment severity and jurisdiction characteristics. Moreover, conditional on the characteristics we consider in Section 4.2—population density, in particular—we find no relationship between punishment severity and the charge to crime ratio.

To evaluate selection into specific arrest charge, we replace the granular arrest charges used to control for underlying conduct in our baseline regression models with a *coarse* measure of initial court charges. The motivation for using a coarse charge type is that, conditional on underlying criminal conduct that leads a charge to be filed, police and prosecutors have little discretion over whether the charges filed are categorized as violent, property, drug, or other. If unobserved charge upgrading or downgrading is substantively influencing the punishment severity estimates, then we would expect this aggregation to meaningfully change the results. Hence, if using coarse charges does not change punishment severity estimates, this suggests unobserved charging decisions are unlikely to be a relevant source of cross-jurisdiction variation in punishment.

While we have over 400 types of court charges across our states, for our coarsened measure, we group offenses into four categories: property, violent, drug, and other. In Table 6, we correlate our original punishment severity estimates with punishment severity estimates derived using coarsened arrest charges. Across states, this correlation ranges from 0.98 to 0.99. Thus, while the mapping of underlying conduct to specific arrest charge may vary across jurisdictions, this distinction is unlikely to bias our punishment severity estimates.

3.1.3 Charges versus Cases

Although we conduct our baseline analysis at the charge level rather than the case level for simplicity, this may introduce bias if co-charges contribute to charge outcomes and charge composition

[18] Summary statistics for the charge to crime ratio are reported in Appendix Table A1.

within cases varies by jurisdictions.

In case-level specifications, we redefine y_{ict} as an indicator for whether a case results in any confinement sentence. Rather than control for arrest charge interacted with criminal history and arrest year ($\tau_{cth(i,t)}$), we control for both the most severe arrest charge in the case and the number of additional misdemeanor and felony charges in the case, interacted with criminal history and arrest year. We also look at cases that consist of only a *single charge*, where there is no distinction between charge and case.

The coefficient estimates for case-level and single charge versions of equation (1) are presented in Appendix Table A2 and Appendix Table A3. We also correlate our baseline punishment severity estimates with case-level and single charge analogs in Appendix Table A4. Estimates are very similar across approaches, with correlations ranging from 0.89 to 0.99.

3.1.4 Using Alternative Charge Outcomes

We next examine alternative measures of punishment severity based on two different charge outcomes: whether the charge results in a conviction, and the sentence length associated with the charge. We again estimate (1) separately by state, but replace the outcome variable. Given the skewed distribution of sentence length and the frequency of zero values, we use two transformations of sentence length as outcomes: an indicator for a sentence of at least 90 days, and an inverse hyperbolic sine (asinh) transformation of sentence length.[19] We measure sentence length in days. For charges that do not result in a jail or prison sentence, we record the sentence length as zero.

A key advantage of our data is that they include charges that are dropped or result in no incarceration sentence. By comparison, many studies use data that only include convictions or charges that lead to incarceration sentences. These more limited data can lead to misleading conclusions about the relative punishment severity of jurisdictions if the conviction or incarceration margin is an important source of variation across jurisdictions. For the sake of comparison, we also include a punishment severity measure derived using the inverse hyperbolic sine of sentence length, but limited to charges that result in *any incarceration sentence*.

Correlations between severity measures are presented separately by state in Panel B of Table 6. The correlation between confinement- and conviction-based severity measures ranges from 0.51 to 0.64 across states. Outside of Texas, the correlation between the confinement-based severity measure and measures derived from sentence length are highly correlated: for the measure based on sentences that are at least 90 days, they range from 0.72 to 0.83; for the measure based on transformed sentence length, they range from 0.95 to 0.97. In Texas, the correlations are more modest: 0.38 for the measure based on sentences that are at least 90 days, and 0.60 for the measure

[19] The asinh function closely parallels the natural logarithm function, but is well defined at zero (Card and Dellavigna, 2019).

based on transformed sentence length.

In general, there is a strong correlation between these measures, where the correlation is stronger between the baseline confinement-based measure and sentence length-based measures. Moreover, when we examine jurisdiction characteristics that correlate with severity in Section 4, the patterns we identify are qualitatively similar across severity measures.

By contrast, our conditional sentence length measure is weakly and *negatively* correlated with the baseline confinement-based measure. Without data on charges that do not lead to an incarceration sentence, we would substantively mischaracterize punishment severity by jurisdiction. This illustrates the importance of using data that includes charges that are dropped or result in no incarceration sentence.

3.2 Validating Estimates Using 'Mover' Defendants

In the analysis above we control for rich offense and charge characteristics that should account for a substantial portion of factors other than jurisdiction-specific punishment severity that determine charge outcomes. However, it is possible that there are critical unobservable determinants that vary across jurisdictions. For example, we do not have direct measures of defendant socioeconomic status, which may affect outcomes directly or through defense attorney quality. We may also miss unobservable severity of the offense or other characteristics of the defendant (e.g. perceived crime risk) that may have important implications for charge outcomes. If these unobservables vary across jurisdictions, they will bias our estimates of punishment severity.

We test for whether unobservables bias our punishment severity estimates by exploiting the fact that many defendants are arrested multiple times and in *multiple jurisdictions.* We use the movement of a defendant from one jurisdiction to another as a quasi-experiment for validating our benchmark punishment severity estimates. Within-defendant comparisons net out time invariant defendant characteristics that contribute to charge outcomes, and we can assess the importance of time-varying unobservable factors by exploiting the timing of the defendant's 'move' from one jurisdiction to another.[20] If our benchmark punishment severity estimates are unbiased measures of the causal effect of jurisdiction, then those estimates should provide unbiased predictions for *changes* in confinement rates for a given defendant that moves from one jurisdiction to another. Our approach is inspired by methods developed in the teacher value-added (Chetty et al., 2014), worker-firm wage decomposition (Abowd et al., 1999; Card et al., 2013, 2016), and health care spending (Finkelstein et al., 2016) literatures. Our approach is most similar to Chetty et al. (2014), who validate benchmark measures of teacher value-added using teachers moving from one school to

[20] When we refer to defendants 'moving' from one jurisdiction to another, we are referring to changes in the jurisdiction where they are arrested, not necessarily changes in residence.

another as quasi-experiments. To the best of our knowledge, this is the first mover-based empirical strategy applied in the criminal justice literature.

In Appendix Table A5, we compare charge and individual characteristics for 'mover' defendants, those who have been arrested in multiple jurisdictions, versus 'stayer' defendants, those who have only been arrested in one jurisdiction.[21] Among stayer defendants, we also look separately at defendants who have faced multiple cases. Twenty-six percent to 40% of defendants have multiple cases in our data, accounting for 58% to 75% of charges. Among defendants with multiple cases, 26% to 41% are arrested in multiple jurisdictions, accounting for 19% to 33% of all charges. Movers are more likely to face confinement sentences than all stayers, and more likely to face confinement sentences than stayers with multiple cases in all states but Texas. They are less likely to be black than all stayers and stayers with multiple cases.

For mover defendants and stayer defendants with multiple cases, we also compare pre- and post-move case pairs for movers and sequential pairs of cases for stayers in Appendix Table A6, focusing on the main charge. For 37.5% to 50.4% of mover pairs, the main charge is of the same crime type in each case. This range is 40.9% to 69.3% for stayer pairs. For movers, 53.5% to 68.7% of post-move cases are in counties adjacent to the pre-move case.

To implement our mover-based test, we use a split-sample procedure. We first randomly partition defendants in each state into 10 equal-sized subsets. For each subset, we estimate equation (1) using the other 9 subsets. To avoid overfitting, we use these (subset-specific) estimates to predict confinement outcomes for mover defendants in the selected subset. For these mover defendants we compare the actual change in the confinement rate before and after the move to the predicted change, adjusting for offense and criminal history. That is, for a defendant who faces one charge in county A and one charge in county B, we compare the predicted difference in outcomes between the two charges to the actual difference in outcomes. For a regression of the actual difference in outcomes on the predicted difference, a slope coefficient of one would indicate that the punishment severity estimates are unbiased. For more details on estimation, see Appendix C.

In Panel A of Figure 2, we plot these actual changes against predicted changes separately by state, pooling by origin and destination punishment severity quartile.[22] The data points fall roughly on the 45° line. We estimate a slope of 1.00 and intercept of 0.01. We cannot formally reject the null hypothesis that punishment severity estimates are unbiased. We also cannot reject *symmetry* for moves to more punitive and less punitive jurisdictions.[23]

This finding has two important implications. First, we can predict within-defendant changes in

21 We limit movers to defendants whose pre-move offense occurs at least two years prior to the end of the data to avoid selecting on initial sentence length.

22 This follows an analogous specification check developed in Card et al. (2016).

23 In particular, if we fit a two-piece linear spline with the knot set at zero, we cannot reject that the two slopes are equal.

confinement remarkably well using data on all defendants. This indicates that punishment severity estimates for all defendants are similar to punishment severity estimates for movers. Second, these predictions are accurate for a variety of defendants as defined by their origin and destination jurisdictions.

The identifying assumption that underpins this validation strategy is that 'mover' defendants do not sort across jurisdictions in a manner that relates to: (1) time-varying unobservable defendant-level or jurisdiction-level determinants of charge outcomes or; (2) match effects—interactions between punishment severity and defendant characteristics.

For example, if defendants that move to a particular jurisdiction are also committing increasingly (and unobservably) more severe crimes, then we would mistakenly identify the jurisdiction as punitive. If a jurisdiction is particularly lenient for drug cases but not other cases, and defendants are more likely to commit drug crimes in that jurisdiction, then we would mistakenly identify this jurisdiction as lenient, when in fact it is only lenient for a particular type of case.[24] We assess these two assumptions in the next section.

3.2.1 Do Defendants Sort on Time-Varying Unobservables or Match Effects?

First, we test whether defendants sort on time-varying unobservables using a placebo test adapted from Card et al. (2013). In particular, we test for pre-trends in mover defendant confinement rates prior to the defendant's change in jurisdictions. To do this, we focus on confinement rates for defendants that are charged in multiple cases in one jurisdiction, and subsequently in at least one case in a different jurisdiction.[25] As an illustrative example, consider a defendant that faces criminal cases 1 and 2 in county A, and criminal case 3 in county B. If defendants are sorting on time-varying unobservables, we may see pre-trends in punishment *prior* to the defendant's move, conditional on observable case and defendant characteristics. To test for such pre-trends, we can thus check whether the identity of county B predicts the difference in outcomes between cases 1 and 2.[26] If sorting on time-varying unobservables is not a factor, then future moves should not predict changes in outcomes between cases 1 and 2.

In Panel B of Figure 2, we plot within-jurisdiction changes in confinement rates against predicted changes based on future moves. The points roughly fall on the horizontal line at zero, and we cannot formally reject the null hypothesis that the slope is zero. This indicates that future moves

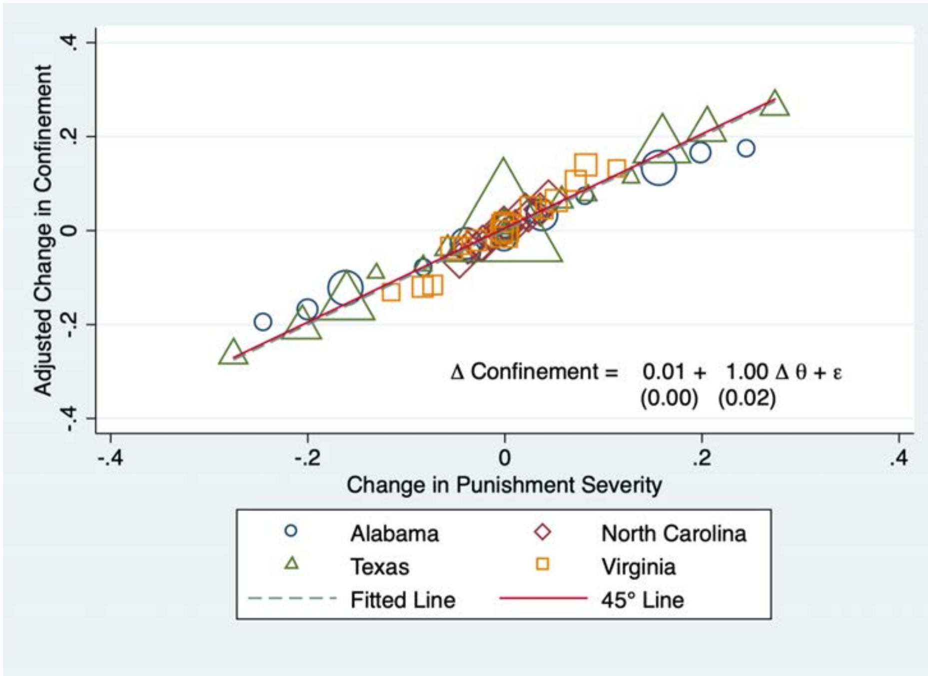
24 There may also be match effects that are specific to movers. For example, some jurisdictions may be more punitive with 'out of town' defendants than long-term residents. However, if punishment severity estimates predict mover confinement rates well, this would imply this type of match effect is not important empirically.

25 In Appendix Figure A1, we replicate Figure 2 Panel A for this sample of defendants.

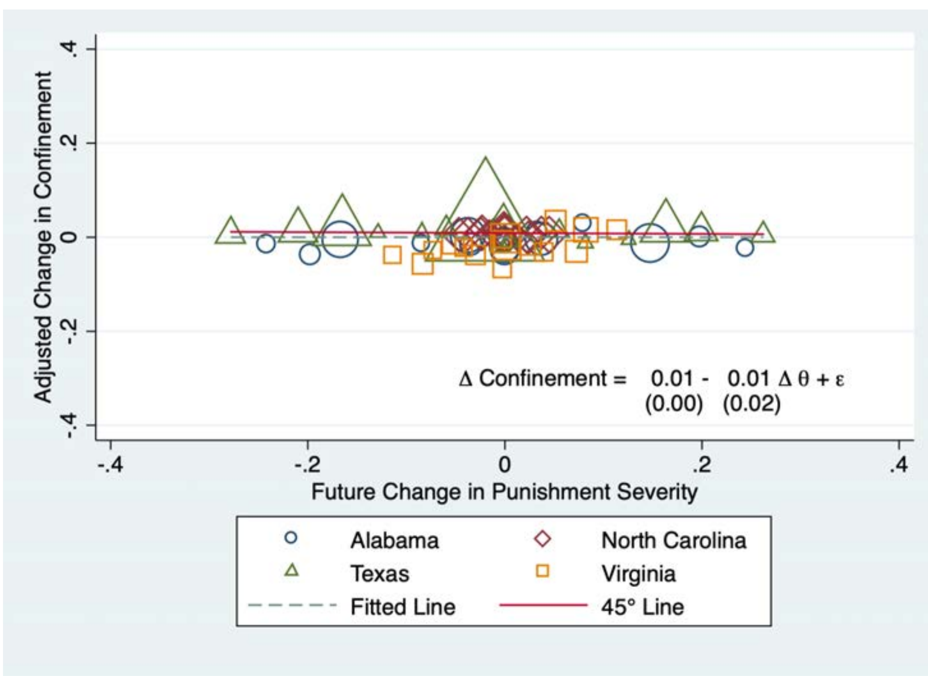
26 Sorting across jurisdictions based on time-varying unobservables would introduce bias, for example, if defendants that committed increasingly (unobservably) serious crimes were also more likely to relocate to less punitive locations.

Figure 2: Mover Defendant Event Studies

(a) Movers



(b) Placebo



Note: In Panel A we plot adjusted realized changes in confinement rate before and after the move against predicted changes by state, adjusting for offense and criminal history, and pooling by origin and destination punishment severity quartile. Marker size is proportional to the number of charges represented in the origin quartile by destination quartile by state cell. The dashed line is the $45°$ line, while the solid line is a fitted line through the points, weighted by cell size. In Panel B we plot within-jurisdiction changes in confinement rates against predicted changes based on future moves. The dashed line is a horizontal line overlapping with the horizontal axis, while the solid line is a fitted line through the points, weighted by cell size.

do not predict earlier changes in confinement rates.

Second, in Appendix C we test whether mover defendants sort on match effects across jurisdictions. We find that they do not, at least on the basis of jurisdiction by crime type or jurisdiction by criminal history match effects. We also documented in Section 3.1.1 that the scope for match effects appears to be limited.

3.2.2 Decomposing Punishment Severity

In this section we quantify the role of punishment severity in explaining cross-jurisdiction variation in confinement rates by decomposing that variation into several components. Our approach follows Finkelstein et al. (2016).

To decompose cross-jurisdiction variation in confinement rates, we first estimate a variant of equation (1) that includes defendant fixed effects, separately by state:

$$y_{ict} = \tau_{cth(i,t)} + \gamma_i + \theta_{j(i,c,t)} + \epsilon_{ict} \tag{2}$$

where γ_i are defendant fixed effects. Note that jurisdiction effects $\theta_{j(i,c,t)}$ in this model are only identified within a connected set (Card et al., 2013). Fortunately, within each state the connected set includes all jurisdictions.

We summarize punishment severity estimates derived from equation (2) in Panel B of Table 7. Consistent with Figure 2 Panel A, these estimates are very similar to the benchmark estimates. The correlation between estimates within states ranges from 0.82 in Alabama to 0.96 in Texas.[27]

In Panel B of Table 7 we present an additive decomposition of the difference between the top quartile and bottom quartile jurisdictions by confinement rate, separately by state. To define the decomposition formally, let $\bar{y}_j$ denote the expectation of y_{ict} across defendants and charges in jurisdiction j. Let $\bar{\gamma}_j$ and $\bar{\tau}_j$ denote the expectation of γ_i and $\tau_{cth(i,t)}$ across defendants and charges in jurisdiction j. Then the difference in confinement rates between two jurisdictions is the sum of punishment severity, defendant, and charge components:

$$\bar{y}_j - \bar{y}_{j'} = \underbrace{(\theta_j - \theta_{j'})}_{\text{jurisdiction component}} + \underbrace{(\bar{\gamma}_j - \bar{\gamma}_{j'})}_{\text{defendant component}} + \underbrace{(\bar{\tau}_j - \bar{\tau}_{j'})}_{\text{charge component}} .$$

We define the share of the difference between two areas attributable to a given component as the ratio of the component difference and the overall difference in confinement rates. When referring to quartile Q_i this is comprised of multiple jurisdictions. We abuse notation and let $\bar{y}_{Q_i}$, θ_{Q_i}, $\bar{\gamma}_{Q_i}$, and $\bar{\tau}_{Q_i}$ denote the simple averages of $\bar{y}_j$, θ_j, $\bar{\gamma}_j$, and $\bar{\tau}_j$ across jurisdictions in Q_i.

[27] The variation is slightly larger for estimates using defendant fixed effects, due at least in part to added measurement error.

Table 7: Summary of Punishment Severity Estimates: Overall vs. Within-Defendant

	Alabama	North Carolina	Texas	Virginia
Avg. Confinement Rate (%)	18.9	7.7	23.6	19.1
σ (Overall)	10.7	2.0	11.2	4.8
σ (Defendant FE)	11.1	2.5	11.8	5.9
Correlation	0.82	0.89	0.96	0.91
Decomposition: Q1 versus Q4 by punishment severity				
Difference in confinement rate				
Overall	28.3	5.8	30.3	14.2
Jurisdiction	22.7	5.4	27.9	12.4
Defendants	5.3	0.7	2.8	1.2
Charges	0.3	-0.3	-0.5	0.6
Share (%) of difference due to				
Jurisdiction	80.2	93.1	92.1	87.3
Defendants	18.7	12.1	9.2	8.5
Charges	1.1	-5.2	-1.7	4.2
Number of Jurisdictions	67	100	253	118

Notes: The top panel compares punishment severity estimates derived with and without defendant fixed effects. 'Overall' punishment severity estimates are derived by estimating equation (1) separately by state and then adding a state-specific constant as described in Section 3. 'Defendant FE' punishment severity estimates are derived by estimating equation (2) separately by state and then adding a state-specific constant. The bottom panel decomposes differences in confinement rates between the top and bottom quartile jurisdictions by punishment severity ($Q1$ and $Q4$), separately by state. The first row reports the difference in average confinement rates between the two sets of jurisdictions ($\hat{Y}_{Q1} - \hat{Y}_{Q4}$); the second row reports the difference due to jurisdiction ($\hat{\theta}_{Q1} - \hat{\theta}_{Q4}$); the third row reports the difference due to defendants ($\hat{\gamma}_{Q1} - \hat{\gamma}_{Q4}$); the fourth row reports the difference due to charge and defendant criminal history ($\hat{\tau}_{Q1} - \hat{\tau}_{Q4}$). The next three rows report the share of the difference in confinement rates due to jurisdiction, defendants, and charge and criminal history.

In Panel B of Table 7 we report the results of this additive decomposition, separately by state. We find that jurisdiction effects explain the bulk of the difference, ranging from 80.2% in Alabama to 93.1% in North Carolina.

In Appendix C, we also decompose cross-jurisdiction variation of confinement rates into the variances and covariances of θ_j, $\bar{\gamma}_j$, and $\bar{\tau}_j$. We find that if jurisdiction effects θ_j were equalized across jurisdictions, cross-jurisdiction variation in confinement rates would be reduced by 64%-93%.

4 Racial Divisions and Punishment Severity

We have provided evidence in support of a causal interpretation of our punishment severity estimates and established the robustness of our severity measure across alternative outcomes and approaches. We next identify jurisdiction-level characteristics that predict punishment severity. To guide this analysis, we sketch a simple model of preferences for punishment based on racial ingroup bias to derive a predicted relationship between punishment severity and local racial heterogeneity.

4.1 A Simple Model

For the purposes of our model, we assume that local residents have to choose an optimal level of punishment, but are constrained to choose an overall punishment severity rather than separate punishment severities by race.[28] Given this restriction, we model the utility of individual i as follows:

$$u_i(s; p(r_i)) = s \times [\alpha(1 - p(r_i)) + \beta p(r_i)] - c(s)$$

where r_i is the racial group of individual i, $p(r_i)$ is the probability that an offender arrested in individual i's home jurisdiction is a member of individual i's racial group, and $c(s)$ is a strictly increasing and convex function (with $c(0) = 0$) characterizing the fiscal and non-pecuniary costs associated with higher severity s.[29] In the expression for individual utility, α and β reflect the relative utility gains associated with punishing outgroup members versus punishing ingroup members (i.e. a negative-valued β implies disutility associated with punishing ingroup members). Based on

[28] This assumption is justified empirically by the findings that (1) incarceration policy severity in a given jurisdiction is highly correlated across racial groups and (2) there is no consistent relationship in our sample between those jurisdiction characteristics that predict overall jurisdiction-level severity and the gap between within-jurisdiction black and white defendant-specific severity parameters. The latter finding is discussed in more depth below.

[29] For example, increased punishment s may impose an additional non-pecuniary cost to the extent that an increase in the likelihood of type II errors, whereby innocent individuals are incorrectly punished, decreases utility (due either to fairness concerns or an individual's self-interested concern that he/she may be erroneously convicted of a crime).

the existing literature related to racial group ingroup bias, we make the assumptions that $\alpha > 0$ and $\alpha > \beta$.[30]

To characterize how predicted punishment preferences vary as a function of local racial composition, first consider a jurisdiction in which a substantial majority of offenders are white (i.e., $p_w >> \frac{1}{2}$). In this case, the punishment severity preferred by white residents, $c'^{-1}(\alpha(1-p_w)+\beta p_w)$, will be lower than $c'^{-1}(\alpha(p_w)+\beta(1-p_w))$, the punishment severity preferred by black residents. Now, suppose that there is a pivotal (median) voter whose preferences determine the jurisdiction-specific punishment severity. Since racial population shares are highly correlated with the share of defendants of each race, the likelihood that the pivotal voter is white is increasing in the share of defendants that is white, and so white punishment preferences will determine local severity. Next, note that as the black share of offenders $(1-p_w)$ increases, the punishment severity preferred by white residents will also increase given that $\alpha > \beta$ and that $c'^{-1}(\cdot)$ is a strictly increasing function by construction. Hence, the punishment severity chosen by the median voter is increasing in black offender share until the median voter switches from a white to black resident. By the symmetry of the model, the punishment severity preferred by black residents is falling as the black share of offenders continues to rise. Consequently, the model predicts that local punishment severity as a function of the black share of offenders will follow an inverted U-shape.

4.2 Testing the Model

Our model predicts a particular non-monotonic causal relationship between local racial composition and punishment severity. To test the model, we would ideally identify a source of exogenous variation in racial composition across jurisdictions, and use that variation to test whether the causal relationship between racial composition and punishment severity exhibits the inverse U-shape pattern the model predicts. Unfortunately, we are unaware of any natural experiment that would provide suitable variation. Instead, we test for an inverted U-shape pattern in the cross-section and adjust for other covariates. An important concern with this approach is omitted variable bias—unobserved differences across jurisdictions may drive any observed relationship between racial composition and punishment severity. Despite this, we believe our 'selection on observables' test is compelling, particularly due to the specific inverse U-shape pattern we are testing for. As we will argue, it is not clear what alternative explanation would be consistent with this pattern.

As an initial test of the prediction derived from the model, Panels A and B of Figure 3 plot transformed punishment severity for each county as a function of its racial composition. To mea-

[30] Luttmer (2001) and Chen and Li (2009) provide observational and experimental support for these assumptions. Anwar et al. (2012) finds that all-white jury pools convict black defendants significantly more often than white defendants, and this gap in conviction rates is eliminated when the jury pool includes at least one black member. These findings are consistent with jurors preferring to punish outgroup defendants over ingroup defendants.

sure racial composition, we use both the black share of the population in 2000 (Panel A) and the black share defendants in that county (Panel B).[31] To make our punishment severity measure comparable across states in this analysis, we transform the measure and express it in terms relative to each jurisdiction's state average. We begin with punishment severity estimates derived from equation (1) using the full data. We then divide this predicted confinement rate by the same severity measure averaged across jurisdictions within the state and take the log of this ratio.[32] The transformed measure is approximately the proportional difference in confinement rates between a jurisdiction and the average jurisdiction in a state, holding other charge characteristics fixed. Given cross-state differences in average predicted confinement rates, we study proportional differences to facilitate cross-state comparisons. Below we denote this transformed punishment severity by $\log \theta'_j$ and refer to this measure as log *relative punishment severity*.

The plot reveals that the inverted U-shaped relationship predicted by our model is indeed borne out in the data. For an initial range of values for the black share of the population or defendants, punishment severity is increasing in the black share. After this range, the sign of the relationship flips.[33]

To clarify this relationship, we pool jurisdictions into bins based on the black share of the population or defendants, where each bin has a range of about 3 percentage points (0-3.3%, 3.3-6.6%, 6.6-9.9%, and so on). For each bin we then average the adjusted punishment severity measure, $\log \theta'_j$. The results are presented in Panels C and D of Figure 3. Marker sizes are proportional to the number of jurisdictions represented in a bin. Note that, for clarity, the span of the vertical axes is substantially narrower in these panels. There is a clear non-monotonic relationship between the black share of the population or defendants and punishment severity, where punishment severity is initially increasing in black share and then the sign of the relationship flips. We use regression models below to measure the implied 'peak' value for the black share.

Note that, if population and defendant shares are equal, voting rates are uniform, voters have uni-dimensional preferences that are homogeneous by race, and all voters are either white or black, the model predicts a peak where the black share of offenders/population is equal to one half. In practice, it is not surprising that we find a peak where the black share of the population is below 0.5. Existing research documents less punitive preferences among blacks than whites (Bobo and

[31] In the model, individual preferences depend on the racial composition of offenders, but the identify of the pivotal voter depends on the composition of the electorate. In practice, the black share of defendants and the black share of the population are highly correlated.

[32] Since regression models include state fixed effects, this normalization does not alter regression results but facilitates data visualization by demeaning logged values separately by state.

[33] Since punishment severity is estimated with controls for defendant demographics, including race, comparisons across jurisdictions reflect a weighted average of differences in the severity of treatment of black and white offenders (with weights determined by jurisdiction-specific offender shares). This approach eliminates the mechanical relationship between local severity and local black defendant share that would otherwise bias cross-jurisdictional comparisons.

Johnson, 2004). Then, to the extent that there is preference heterogeneity such that some white residents have less punitive preferences and do not exhibit ingroup bias, we should anticipate a peak below 0.5.[34]

We next move to a more thorough analysis of the relationship between local punishment severity and racial composition. Absent any source of plausibly exogenous cross-sectional variation in racial composition, we introduce a series of additional jurisdiction-level covariates into a regression of log adjusted punishment severity on a quadratic in the black share of the population to assess the extent to which alternative mechanisms may drive the observed relationship. Specifically, we estimate models of the following form:

$$log\ \theta_j^{'} = x_j\beta + \tau_s + \epsilon_j \tag{3}$$

where $log\ \theta_j^{'}$ is the log adjusted punishment severity described above, x_j is a vector of jurisdiction characteristics, and τ_s is a set of state fixed effects.

Researchers studying US trends in crime and punishment have highlighted the important role that historical violent crime rates played in driving the increased severity of punishment over recent decades and in generating cross-state variation in punishment severity (see, for instance, Western, 2006). To test whether local variation in past crime rates is associated with differences in punishment severity *within* states, we control for measures of growth in violent crime rates between 1970 and 1990 and the 2000 violent crime rate, both measured at the jurisdiction level.[35] Each measure is standardized to have a mean of zero and a standard deviation of one. The crime measures are derived from FBI UCR data. In addition to a quadratic in the black share of the population or defendants, we include log average household income, the Gini index of income inequality, the fraction of prime-aged males in the population, and log population density, all measured in 2000. Descriptive statistics for these county characteristics are reported in Appendix Table A7. There is one observation per jurisdiction. Note that we are missing data on crime and the Gini index for some counties.[36] In the regression models, we set missing values to zero and include indicators for

[34] Differences in voting rates by race, in the share of the population categorized as "Other race", and the multi-dimensionality of policy preferences would also generate uncertainty in the precise level of the black population share at which punishment severity peaks. Although Republican Party support is undoubtedly an imperfect proxy for punishment preferences and does not capture preference intensity, a back-of-the-envelope calculation relying on race-specific party affiliation reveals that the black population share at which we would expect to observe the median voter change from Republican to Democrat ranges from 0.23 in North Carolina to 0.40 in Alabama.

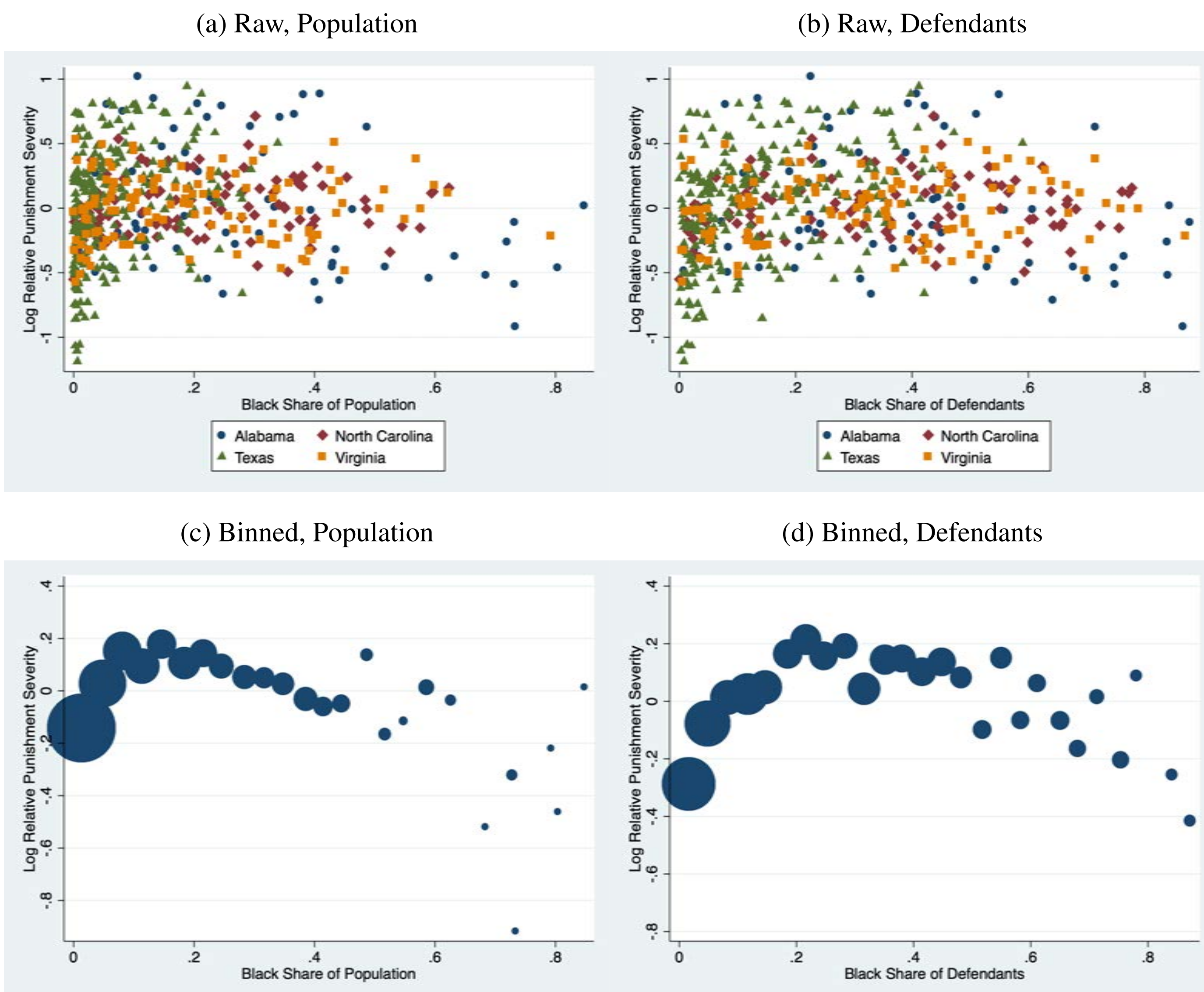
[35] We calculate the growth in violent crime as

$$r_{\text{growth}} = \frac{r_{1990} - r_{1970}}{0.5r_{1990} + 0.5r_{1970}}$$

where r_{1990} and r_{1970} are the local violent crime rates in 1990 and 1970.

[36] We are missing data on the violent crime rate in 2000 for seven counties, on violent crime rate growth from 1970 to 1990 for eight counties, and the Gini index for 67 counties.

Figure 3: Punishment Severity and Racial Heterogeneity



Note: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. In binned plots, marker sizes are proportional to number of jurisdictions represented in bin.

missing data for each of these covariates as additional controls.

Regression estimates are presented in Table 8. In columns (1)-(5) we use the black share of the population and its square to measure a jurisdiction's racial composition. In columns (6)-(10), we use the black share of defendants and its square. The results are similar for both measures. We discuss the results using the black share of the population first, and then discuss the differences in results between the two measures.

Column (1) presents the regression equivalent of Figure 3, with no controls other than the black share of the population, its square, and state fixed effects. Point estimates are consistent with an inverted U-shaped relationship between local severity and black share of the population and imply that punishment severity is highest in jurisdictions with a black share of the population equal to 0.3. At this maximum, the predicted value of θ is 24 log points larger than the predicted value where black share is set to zero. This implies that predicted punishment severity is 27% higher in jurisdictions with this level of heterogeneity relative to all-white jurisdictions.[37]

Columns (2) and (3) add our set of jurisdiction-level controls: log population density, log average household income, the Gini index of income inequality, and the fraction of prime-aged males in the population. The only difference between the two models is the measure of local crime that we include as a control. Column (2) uses the growth in the violent crime rate from 1970 to 1990, and column (3) uses the violent crime rate in 2000.[38] In both specifications, the inverted U-shape relationship between punishment severity and black population share remains highly significant, though is somewhat muted in magnitude. The peak value for the black share of the population moves up to 0.33 in column (2) and to 0.37 in column (3). At these peak values for columns (2) and (3), the predicted value of θ is 14 and 17 log points larger than the predicted value where black share is set to zero, respectively. The coefficient on the growth in violent crime in column (2) is close to zero and statistically insignificant, while the coefficient on violent crime in 2000 in column (3) is *negative* and small in magnitude, though statistically significant. Results from these specifications lend little support to the hypothesis that within-state variation in present-day severity is explained by historical crime waves or current crime patterns. Turning to the remaining covariates, population density also consistently predicts higher confinement rates. A jurisdiction with 10% higher population density is predicted to be about 1% more punitive.

Given that population density is a strong predictor of severity, one concern is that the relationship we identify between racial composition and punishment severity is driven in part by a nonlinear relationship between population density and severity. Column (4) repeats the specifica-

37 As an alternative approach to testing for an inverted U-shaped relationship between black population share and punishment severity, we estimate two piece linear splines and test for a positive initial slope and negative final slope. If we set the knot point to 0.3, we estimate an initial slope of 0.928 (standard error 0.184) and final slope of -1.151 (0.224).

38 Results are similar if we use total Part I crime rates rather than restricting to violent crime.

Table 8: Punishment Severity and Racial Heterogeneity

Outcome:	Log Relative Punishment Severity									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black Share of Population	1.644**	0.853**	0.946**	0.858**	0.895**					
	(0.297)	(0.257)	(0.257)	(0.257)	(0.279)					
Black Share of Population, Squared	-2.763**	-1.282**	-1.279**	-1.013**	-1.494**					
	(0.501)	(0.405)	(0.396)	(0.383)	(0.450)					
Black Share of Defendants						2.147**	1.298**	1.336**	1.177**	1.424**
						(0.270)	(0.255)	(0.255)	(0.264)	(0.279)
Black Share of Defendants, Squared						-2.705**	-1.725**	-1.682**	-1.375**	-1.905**
						(0.353)	(0.342)	(0.342)	(0.347)	(0.389)
Log Population Density		0.085**	0.099**	†	x		0.081**	0.092**	†	x
		(0.020)	(0.021)				(0.020)	(0.021)		
Log Average Household Income		0.394*	0.362*	0.344*	x		0.295~	0.259	0.255	x
		(0.163)	(0.164)	(0.171)			(0.157)	(0.160)	(0.168)	
Gini Coefficient		-0.194	-0.070	-0.011	x		-0.055	0.050	0.076	x
		(0.220)	(0.226)	(0.228)			(0.219)	(0.226)	(0.225)	
Fraction Males Aged 15-29		0.077	0.054	0.233	x		0.009	-0.011	0.135	x
		(0.744)	(0.732)	(0.741)			(0.732)	(0.719)	(0.734)	
Violent Crime Rate Growth, 1970-1990		-0.017		-0.027~			-0.017		-0.024	
		(0.016)		(0.016)			(0.016)		(0.015)	
Violent Crime Rate, 2000			-0.044*	-0.042*	x			-0.038*	-0.038*	x
			(0.018)	(0.017)				(0.017)	(0.017)	
State FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.049	0.202	0.206	0.222	0.238	0.111	0.227	0.230	0.239	0.266
Observations	538	538	538	538	538	538	538	538	538	538

Notes: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls.
'†' denotes inclusion of a five-piece linear spline in log population density. 'x' denotes inclusion of the covariate interacted with state fixed effects.
Robust standard errors in parentheses. ~ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

tion in column (3) but adds a 5-piece linear spline in log population density as controls. Controlling for population density in this more flexible manner has little effect on the coefficient estimates for the black share of the population and its square.

In column (5) we allow each of the non-race covariates to vary by state, interacting each with state indicator variables. The inverted U-shape relationship between punishment severity and black population share remains highly significant and unchanged in this specification that controls more flexibly for the full set of non-race covariates.

The pattern of coefficients is similar in columns (6)-(10) where we replace the black share of the population with the black share of defendants. There are two noticeable differences. First, the implied peak moves to about 0.4. Second, the difference between predicted θ at 'peak' heterogeneous jurisdictions and all-white jurisdictions increases to 43 log points without additional controls and to 24 log points with controls. Both findings are consistent with what we see graphically in Figure 3.

4.2.1 Robustness Checks and Extensions

In this section we explore the robustness of our results along a number of dimensions.

First, we examine whether the relationship between punishment severity and racial composition that we identify is present for our alternative measures of severity based on conviction rates and sentence length. We estimate equation (3) but replace the outcome used to measure punishment severity. The results are presented in Table 9. In columns (1) and (2) the outcome is conviction. In columns (3) and (4) the outcome is a sentence above 90 days. In columns (5) and (6) the outcome is inverse hyperbolic sine-transformed sentence length. In odd columns we use the black share of the population and its square to measure a jurisdiction's racial composition, while in even columns we use the black share of defendants and its square. Across outcomes, we see a similar inverted U-shape relationship between punishment severity and black share that peaks for black share values in the 0.3 to 0.4 range.

In columns (7) and (8) the outcome is inverse hyperbolic sine-transformed sentence length, but limited to charges with any incarceration sentence. We include this measure to see what we would have concluded if our data excluded dropped charges and those not leading to an incarceration sentence. Strikingly, we see little to no relationship between this measure and the black share of the population or defendants.

Second, we address the concern raised in Section 3.1.2 that the type of offenses that lead to charges may vary across counties. For example, jurisdictions with fewer marginal charges may appear more severe in part because the composition of offenses that actually lead to a charge may be (unobservably) more serious. We estimate versions of equation (3) that include a jurisdiction's charge to crime ratio as an additional control. To match the coverage of the UCR crime data, we also replace the baseline punishment severity measure with a measure derived from only violent and property crimes in some specifications. The results are presented in Appendix Table A8. We find that, conditional on the jurisdiction covariates we include, the charge to crime ratio is uncorrelated with punishment severity and its inclusion has no effect on the coefficients for black share.

Third, we address the concern that the inverted U-shape relationship identified in Table 8 can be explained by endogenous migratory responses to local punishment severity or to other correlated community characteristics. In Appendix Table A9 we replace the black share of the population measure with the 1860 county-level share of the population that was enslaved. Despite the fact that historical data is only available for two-thirds of the jurisdictions in the sample, we identify a similarly robust inverted U-shape relationship between 1860 slave share and contemporaneous punishment severity.

Fourth, to assess the validity of the assumption that jurisdiction residents' preferences determine average local punishment severity rather than *race-specific* punishment severity, we re-estimate the specifications included in Table 8 in Appendix Table A10 but use the black-white difference in log adjusted local severity as our outcome measure. While the coefficients on black population share and its square are statistically significant in more sparse specifications, these coefficients are no longer statistically significant when we allow for state-specific slopes for non-race jurisdiction characteristics. When we measure black share using the composition of defendants, the coefficients on black share and its square are small in magnitude, statistically insignificant, and of inconsistent sign across specifications. Overall, we do not find robust evidence that race-based gaps follow the same inverted U-shape pattern as overall punishment severity. Moreover, as described below, when we construct separate punishment severity measures for black and white defendants, we find an inverted U-shape pattern for both measures.

Fifth, we examine whether the non-monotonic relationship we identify between punishment severity and black share is present for other jurisdiction characteristics. In Appendix Figure A2, we plot the relationship between each covariate included in Table 8 and punishment severity. To the extent that racial divisions indeed explain the inverted U-shape relationship between punishment severity and black population or defendant share, we should not expect to see a similar non-monotonic relationship between local severity and any of the other included covariates. Reassuringly, there is indeed no evidence of a non-monotonic relationship between any of the other included covariates and punishment severity.

Sixth, we test whether our results are robust to using punishment severity derived from the model (2) that includes defendant fixed effects or subgroup-based estimates explored in Section 3.1.1. In Appendix Table A11 we show the same inverted U-shape relationship is present for each alternative measure of punishment severity.[39]

Seventh, to assess the degree of potential bias due to unobservables, we use the approach outlined in Oster (2019).[40] We show in Appendix Table A13 that selection on unobservables would need to be over two times as large as selection on observables to explain the measured relationship between punishment severity and racial composition. These estimates are notably above the upper bound of one suggested in Oster (2019) for calculating bias-adjusted treatment effects. Moreover, if we include population density in our baseline model, the implied degree of selection on unobservables that would be required to explain our estimates increases to between 3.8 and 12.4 times as large as selection on observables.

An alternative explanation for the relationship we identify between racial heterogeneity and punishment severity is that (1) a higher share of defendants in racially heterogeneous communities are paired with judges or prosecutors of another race and (2) judges or prosecutors treat outgroup members more severely than ingroup members. Given the paucity of black prosecutors, ingroup bias seems unlikely to explain the pattern we observe. In 2014, only 6.6% of chief prosecutors are black in our sample states, and that drops to 2.5% if we exclude Virginia (Reflective Democracy Campaign, 2018). While Shayo and Zussman (2011) documents robust evidence of judicial ingroup bias in Israel, findings from the US are mixed and suggest that ingroup bias among judges may be limited. Cohen and Yang (2019) finds that among Republican-appointed federal judges, white judges differentially punish black defendants more severely. However, the authors do not find differential gaps in punishment among Democratic-appointed judges and note that the vast majority of black federal judges are Democratic-appointed. Schanzenbach (2015) finds that federal judges do not exhibit ingroup bias, and Arnold et al. (2018) finds no evidence that racial bias varies with judge race among bail judges in Philadelphia and Miami-Dade counties. While Abrams et al. (2012) finds that black judges impose relatively short sentences on black defendants, they are not less likely to impose confinement sentences on black defendants. Our own finding that the black-white gap in punishment severity does not vary in a consistent manner with local racial composition also suggests that judicial ingroup bias is unlikely to explain the relationship between racial heterogeneity and overall punishment severity that we identify. If, for instance, white-majority jurisdictions elected white judges who punished black defendants more severely, we should identify a negative relationship between the black share of the population and the black-white gap in local punishment severity.

To provide support for the hypothesis that local racial composition affects punishment severity through the preferences of the local electorate, Appendix Table A14 employs jurisdiction-level data

[39] We also show in Appendix Table A12 that the same inverted U-shape relationship is present when observations are weighted by jurisdiction population.

[40] We use the command *psacalc* in Stata to implement relevant calculations.

on support for statewide ballot measures related to the punishment of criminals and the rights of the accused.[41] We find that increased local support for harsher punishment is strongly associated with higher punishment severity and has the same inverse U-shaped relationship with the black share of the population and with the black share of defendants (though the quadratic term is imprecise when controls are included).[42]

A natural question is whether the relationship that we identify between local racial composition and punishment severity generalizes outside of our sample states. Given that the estimation of jurisdiction-specific punishment severity requires rich defendant- and charge-level data that are not widely-available outside of our sample, it is not feasible to answer this question conclusively. However, we can utilize comparable data from the State Court Processing Statistics Data series, which includes a sample of cases from the nation's largest counties, to make progress in assessing generalizability. Most included states have coverage for two or fewer counties, so we focus on cross-state (as opposed to within-state) analyses. Our findings, presented in Appendix Table A15, reveal an inverse U-shaped relationship between the *state-level* black population share and punishment severity. This pattern is also shown graphically in Appendix Figure A3. Point estimates associated with the county-level black share of the population are comparable to the estimates from Table 8 for the South-only sample, but are imprecise in both the South-only and nationwide samples.[43] The relative magnitudes of these estimates indicate that state-level racial composition may play a stronger role in explaining cross-state variation in punishment severity than local racial composition plays in explaining within-state variation.[44] We speculate that the central role of state-level racial composition may reflect the influence of racial dynamics on state laws, and we hope these suggestive findings motivate future research aimed at better understanding this relationship.

41 Additional details related to the construction of this vote share measure are provided in the Appendix D.

42 Consistent with Cohen and Yang (2019), we also show in Appendix Table A14 that Republican Party support is a strong predictor of punishment severity.

43 In the full-sample specification, we identify a similarly imprecise inverse U-shaped relationship when the explanatory variables characterizing the state-level black share of the population are excluded.

44 Estimates imply that punishment severity is highest in states with a black share of the population equal to 0.17. At this maximum, predicted severity is 82% higher in jurisdictions with this level of heterogeneity relative to all-white jurisdictions. For reference, the measured difference in punishment severity between the most lenient and harshest states in our sample is approximately 250%. In the South-only sample, the peak occurs where the black share of the population is 0.18, though the implied difference in punishment severity between jurisdictions with this level of heterogeneity relative to all-white jurisdictions is much larger (685%). We note, however, that these South-only estimates are based on only nine data points and confidence intervals are wide. Moreover, the most homogeneous Southern state included in the analysis, Kentucky, has a black population share over 7% and so this calculation is particularly reliant on out-of-sample extrapolation.

Table 9: Punishment Severity and Racial Heterogeneity, Alternative Outcomes

Outcome:	Convictions		Sentence $\geq$ 90 Days		asinh(Sentence Length)		asinh(Cond. Sentence Length)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black Share of Population	0.247˜		0.683*		0.973**		-0.049	
	(0.131)		(0.298)		(0.260)		(0.065)	
Black Share of Population, Squared	-0.455*		-1.131*		-1.414**		0.029	
	(0.183)		(0.440)		(0.382)		(0.096)	
Black Share of Defendants		0.375**		0.832**		1.131**		-0.133*
		(0.124)		(0.296)		(0.249)		(0.064)
Black Share of Defendants, Squared		-0.553**		-1.151**		-1.436**		0.134˜
		(0.164)		(0.381)		(0.327)		(0.079)
Log Population Density	0.023*	0.022*	0.085**	0.083**	0.077**	0.073**	0.003	0.004
	(0.009)	(0.009)	(0.021)	(0.021)	(0.019)	(0.019)	(0.004)	(0.004)
Log Average Household Income	0.106˜	0.082	0.096	0.047	0.281˜	0.206	-0.084**	-0.072*
	(0.057)	(0.057)	(0.151)	(0.150)	(0.152)	(0.148)	(0.029)	(0.029)
Gini Coefficient	-0.016	0.023	0.248	0.332	0.200	0.305	0.086	0.079
	(0.147)	(0.149)	(0.293)	(0.298)	(0.233)	(0.237)	(0.053)	(0.053)
Fraction Males Aged 15-29	-0.556˜	-0.562˜	-1.517˜	-1.513*	-0.853	-0.863	-0.144	-0.123
	(0.324)	(0.318)	(0.773)	(0.757)	(0.580)	(0.567)	(0.186)	(0.184)
Violent Crime Rate, 2000	-0.013	-0.011	-0.046*	-0.043*	-0.051*	-0.047*	-0.004	-0.003
	(0.010)	(0.010)	(0.021)	(0.021)	(0.020)	(0.020)	(0.004)	(0.004)
State FEs	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.038	0.048	0.120	0.127	0.143	0.155	0.528	0.531
Observations	537	537	536	536	538	538	538	538

Notes: For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls.
Columns (1) and (2) exclude one jurisdiction (Austin County, Texas) with corresponding punishment severity estimate (in this case, the predicted conviction rate) below zero. Columns (3) and (4) exclude two jurisdictions (Brooks County, Texas and Duval County, Texas) with corresponding punishment severity estimates (predicted rate of sentences $\geq$ 90 days) below zero.
Robust standard errors in parentheses. ˜ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

5 Simulation

Given evidence that there are significant cross-jurisdictional differences in the severity of criminal punishment and that punishment severity can be interpreted causally, we next simulate the share of charges leading to an incarceration sentence and the race-based gap in this share under a counterfactual in which more punitive jurisdictions adopt the punishment level imposed by communities in their state at the tenth percentile of the predicted confinement rate distribution based on black population share. Specifically, within a given state, punishment severities above the predicted level assigned to a jurisdiction at the tenth percentile are adjusted downwards to this level. Table 10 presents a comparison of actual confinement outcomes to the simulated confinement outcomes for whites versus blacks in the four states in our sample.[45] In the simulation, we account for the fact that reduced punishment severity interacts dynamically with our criminal history measures, which are a function of past charge dispositions. In order to do so, we adjust confinement probability to account for the fact that simulated criminal histories will be made shorter than actual criminal histories by the reduction in conviction rates (and confinement rates, in Virginia) imposed. Across all four states in the sample, the magnitude of the race-based confinement gap declines in level terms when we simulate outcomes. Importantly, this is not a mechanical consequence of the adjusted jurisdiction-specific punishment severity. Instead, this finding reflects the fact that black residents of these states disproportionately reside in high-severity jurisdictions. Across states, the black-specific measure of confinement sentences per capita declines by 15-20%, with an average decline of 16%, and the white-specific measure of confinement sentences per capita declines by 17-27%, with an average decline of 19%. Declines in punishment severity correspondingly reduce the magnitude of the gap in confinement sentences per capita by 12-16%, with an average decline of 14%.

6 Conclusion

We study the role that racial divisions play in explaining the punitiveness of US criminal justice policy by collecting and analyzing administrative criminal justice data from four Southern states. We identify substantial variation across jurisdictions within a given state in the severity of punishment and show that this variation persists even when we include a rich set of charge- and defendant-level covariates. We employ a mover-based identification strategy adopted from the teacher value-added and worker-firm wage decomposition literatures and find that unobserved defendant heterogeneity cannot explain the differences in punishment severity that we identify. We

[45] Simulation-based confinement sentences per capita measures do not line up precisely with population-based measures given the additive relationship between defendant covariates and charge dispositions that is assumed when constructing simulated outcomes under alternative counterfactual scenarios.

Table 10: Simulation Results

	Alabama	North Carolina	Texas	Virginia
Confinement Sentences per 100,000				
White (Actual)	592	481	1673	644
Black (Actual)	1595	1658	2555	2065
White (Simulation)	429	391	1396	516
Black (Simulation)	1280	1405	2171	1714
Number of Jurisdictions	67	100	253	118

Notes: Simulated confinement sentences per 100,000 age 15 or above are derived as described in Section 5. Statistics weighted by race-specific jurisdiction population.

proceed to write down a simple model of racial ingroup bias that predicts an inverse U-shaped relationship between local black share of the population and punishment severity. This prediction is borne out in the data. Our analysis concludes with a simulation exercise that shows that punishment levels and race-based incarceration rate gaps would decline by approximately 15%, on average, if more punitive jurisdictions adopted the punishment severity of neighbors that are more racially homogeneous.

While a large literature has documented the connection between racial stratification and support for public goods and redistribution, this research offers novel evidence that racial heterogeneity can be similarly linked to preferences for a 'public bad': more punitive criminal justice policy. Given that blacks are more likely to reside in racially heterogeneous communities in the states in our sample, this finding has important implications for the severity of criminal justice policy faced by the average white versus black resident of these states. Moreover, our findings suggest that large race-based gaps in criminal justice outcomes may persist even in the absence of discriminatory treatment within any given jurisdiction.

References

Abowd, John, Francis Kramarz, and David Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *108.*

Abrams, David, Marianne Bertrand, and Sendhil Mullainathan, "Do Judges Vary in their Treatment of Race?," *The Journal of Legal Studies*, 2012, *41* (2), 347–383.

Alesina, Albert, Reza Baqir, and William Easterly, "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 1999, *114* (4).

Alexander, Michelle, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*, New York: The New Press, 2010.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson, "The Impact of Jury Race in Criminal Trials," *Quarterly Journal of Economics*, May 2012, *127* (2), 1017–1055.

Arnold, David, Will Dobbie, and Crystal S. Yang, "Racial Bias in Bail Decisions," *Quarterly Journal of Economics*, November 2018, *133* (4), 1885–1932.

Berdejó, Carlos and Noam Yuchtman, "Crime, Punishment, and Politics: An Analysis of Political Cycles in Criminal Sentencing," *Review of Economics and Statistics*, July 2013, *95* (3), 741–756.

Bobo, Lawrence D. and Devon Johnson, " A Taste For Punishment: Black and White Americans' Views on the Death Penalty and the War on Drugs," *Du Bois Review*, 2004, *1* (1), 151–180.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *Quarterly Journal of Economics*, 2016, *131* (2).

_ and Stefano Dellavigna, "What Do Editors Maximize? Evidence from Four Leading Economics Journals," *Review of Economics and Statistics,* forthcoming, 2019.

_ , Joerg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128.*

Carson, E. Ann, "Prisoners in 2013," September 2014. U.S. Department of Justice, Office of Justice Programs.

Chen, Yan and Sherry Xin Li, "Group Identity and Social Preferences," *American Economic Review*, 2009, *99* (1).

Chetty, Raj, John Friedman, and Jonah Rockoff, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 2014, *104* (9), 2593–2632.

Cohen, Alma and Crystal S. Yang, "Judicial Politics and Sentencing Decisions," *American Economic Journal: Economic Policy*, February 2019, *11* (1), 160–191.

Collister, Brian, "Texas troopers ticketing Hispanic drivers as white," November 2015.

Dahlberg, Matz, Karin Edmark, and Helene Lundqvist, "Ethnic Diversity and Preferences for Redistribution," *Journal of Political Economy*, 2012, *120* (1).

Dyke, Andrew, "Electoral Cycles in the Administration of Criminal Justice," *Public Choice*, December 2007, *133* (3/4), 417–437.

Enns, Peter K., "The Public's Increasing Punitiveness and Its Influence on Mass Incarceration in the United States," *American Journal of Political Science*, October 2014, *58* (4), 857–872.

Enos, Ryan, "What the Demolition of Public Housing Teaches Us about the Impact of Racial Threat on Political Behavior," *American Journal of Political Science*, 2015, *60* (1).

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *Quarterly Journal of Economics*, November 2016, *141* (4), 1681–1726.

Fischman, Joshua B. and Max M. Schanzenbach, "Racial Disparities under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums," *Journal of Empirical Legal Studies*, 2012, *9.*

Glaser, James, "Back to the Black Belt: Racial Environment and White Racial Attitudes in the South," *Journal of Politics*, 1994, *56* (1).

Goncalves, Felipe and Steve Mello, "A Few Bad Apples? Racial Bias in Policing," January 2018.

Gottschalk, Marie, *Caught: The Prison State and the Lockdown of American Politics*, Princeton, New Jersey: Princeton University Press, 2015.

Hetey, Rebecca C. and Jennifer L. Eberhardt, "Racial Disparitites in Incarceration Increase Acceptance of Punitive Policies," *Psychological Science*, October 2014, *25* (10), 1949–1954.

Huber, Gregory A. and Sanford C. Gordon, "Accountability and Coercion: Is Justice Blind when It Runs for Office?," *American Journal of Political Science*, April 2004, *48* (2), 247–263.

Keen, Bradley and David Jacobs, "Racial threat, partisan politics, and racial disparities in prison admissions: a panel analysis," *Criminology*, 2009, *47*, 209–238.

Keller, Josh and Adam Pearce, "This Small Indiana County Sends More People to Prison than San Francisco and Durham, N.C., Combined. Why?," https://www.nytimes.com/2016/09/02/upshot/new-geography-of-prisons.html= September 2016. The New York Times.

Key, V.O., *Southern Politics in State and Nation*, New York: Knopf, 1949.

Lim, Claire S. H., "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges," *American Economic Review*, June 2013, *103* (4), 1360–1397.

Lim, Claire S.H., James M. Snyder Jr., and David Strömberg, "The Judge, the Politician, and the Press: Newspaper Coverage and Criminal Sentencing across Electoral Systems," *American Economic Journal: Applied Economics*, October 2015, *7* (4), 103–135.

Luttmer, Erzo F.P., "Group Loyalty and the Taste for Redistribution," *Journal of Political Economy*, 2001, *109* (3).

Miethe, T.D., "Charging and Plea Bargaining Practices under Determinate Sentencing," *Journal of Criminal Law and Criminology*, 1987, *78*.

Muhammad, Khalil Gibran, *The Condemnation of Blackness: Race, Crime, and the Making of Modern Urban America*, Cambridge, MA: Harvard University Press, 2010.

Nelson, Michael J., "Responsive Justice? Retention Elections, Prosecutors, and Public Opinion," *Journal of Law and Courts*, Spring 2014, *2* (1), 117–152.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 2019, *37* (2), 187–204.

Pfaff, John, "Escaping from the Standard Story: Why the Conventional Wisdom on Prison Growth is Wrong, and Where We Can Go From Here," *Federal Sentencing Reporter*, April 2014, *26* (4), 265–270.

Raphael, Steven and Sandra V. Rozo, "Racial Disparities in the Acquisition of Juvenile Arrest Records," March 2017.

Reflective Democracy Campaign, "Justice for All? A Project by the Reflective Democracy Campaign on Who Prosecutes in America," May 2018.

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

Schanzenbach, Max M., "Racial Disparities, Judge Characteristics, and Standards of Review in Sentencing," *Journal of Institutional and Theoretical Economics*, March 2015, *171* (1), 27–47.

Shayo, Moses and Asaf Zussman, "Judicial Ingroup Bias in the Shadow of Terrorism," *Quarterly Journal of Economics*, August 2011, *126* (3), 1447–1484.

Unnever, James and Francis Cullen, "Reassessing the Racial Divide in Support of Capital Punishment: The Continuing Significance of Race," *Journal of Research in Crime and Delinquency*, 2007, *44* (1).

Unnever, James D. and Francis T. Cullen, "The Social Sources of Americans' Punitiveness: A Test of Competing Models," *Criminology*, February 2010, *48* (1), 99–129.

Walmsley, Roy, "World Prison Population List: Eleventh Edition," 2016. World Prison Brief.

Western, Bruce, *Punishment and Inequality in America*, New York: Russell Sage Foundation, 2006.

For Online Publication

A Appendix: Data Description

A.1 Alabama

The data for Alabama are from the Alabama Administrative Office of Courts and shared with us by the Center for Science and Law. The earliest records in the data date back to the early 20th century, though data quality and completeness improves over time. We focus on charges filed between 2000 and 2010. We end in 2010 because in the extract we obtained the share of charges that remain unresolved begins to increase significantly in 2011.

We drop charges with missing data on the defendant, including date of birth, gender, and race. In most of these instances, the defendant listed appears to be an organization (e.g., a bail bond company) rather than a person.

We drop charges with missing dispositions, which appear to generally reflect charges that are on-going. We drop probation violations, appeals, and records that indicate intermediate outcomes, such as the transfer of a charge from a lower court to a higher court. We restrict to felony and misdemeanor non-traffic offenses.

To match multiple cases to individuals, we group defendants based on full name and date of birth.

The data include the zip code of the court and a court-specific code, but not the name or the county. We match courts to counties based on the zip code. In ambiguous instances, we manually match charges to counties based on the location of actual courthouses.

In Alabama, criminal cases are handled in Circuit and District Courts. Circuit courts are courts of general jurisdiction, and handle all felony cases. There are 148 Circuit Court judges divided among 41 judicial circuits. District Courts handle misdemeanors. There are 98 judges in 67 District Courts, one court in each county. Each judicial circuit is served by a chief prosecutor ('District Attorney').

Judges for both Circuit and District Courts are elected in partisan elections. The length of term is 6 years. Prosecutors are also elected to 6-year terms in partisan elections. Circuit Court judges and prosecutors are elected at the circuit level. District Court judges are elected at the county level.

A.2 North Carolina

The data for North Carolina are from the North Carolina Administrative Office of Courts. These data contain records for charges initially filed from 2007 to 2014.

To construct the charge-level data file that is ultimately used in our analysis, we merge case records with offense records (that include disposition and sentence outcomes) based on the unique case identifier provided by the North Carolina Administrative Office of Courts, as well as an identifier for the county in which the charge was adjudicated and the case-specific charge number. We successfully merge 100% of charges to disposition records. While the data do not include unique defendant identifiers, we match multiple cases to the same individual using their full name and address.

We restrict the sample to include only offenses classified as felonies or misdemeanors. Next, we exclude charges for which the same charge is subsequently listed with a final disposition. We also drop charges with intermediate outcomes corresponding to the following recorded dispositions: Superseding Indictment or Superseding Process, Transfer to Superior Court, Probable Cause Found, Change of Venue, and Withdrawn from Superior Court. We drop charges with disposition records that contain missing dispositions, since the structure of the data means that charge dispositions should be available for all included charges. We drop charge-level observations corresponding to probation and parole violations, and we drop observations corresponding to youth aged under 16. Finally, we drop observations that are missing information on defendant age.

To construct our confinement and sentence length outcomes, we convert reported incarceration sentence days, months and years into the number of days sentenced. To do so, we rely on the Minimum Sentence Length values associated with each charge disposition. We categorize a charge as resulting in confinement if (1) a non-zero incarceration sentence is listed and no concurrent probation sentence is listed or (2) the charge results in mandatory confinement based on North Carolina structured sentencing guidelines. To identify charge dispositions for charges with missing sentence records, we rely on the offense file and code a charge as resulting in a conviction if the Convicted Offense Code variable is non-missing (i.e., an offense for which the defendant was convicted is provided). To classify charges as dropped, we construct an indicator variable based on whether the disposition is listed as any of the following: Dismissed by the court, Dismissal by DA, No probable cause, Voluntary dismissal DA, Dismissal with leave by DA, and No true bill returned.

Based on guidance received from the North Carolina Administrative Office of Courts, to identify charges corresponding to a single case, we take the connected set of charges that meet any of the following three criteria: (1) charge records include the same case identifier, (2) one charge has a "consolidated for sentencing" case identifier that matches the case identifier associated with another charge, or (3) charges are filed against the same defendant for the same offense code and on the same offense date.

In North Carolina, criminal cases are handled in Superior and District Courts. Superior Courts handle all felony cases. There are 109 Superior Court judges divided among 50 Superior Court districts. These districts are further grouped into 8 divisions. Every 6 months, elected Superior Court judges rotate from one district to another within their division.[46] District Courts handle misdemeanors. There are 256 judges in 47 judicial districts, one court in each county. There are 44 separate prosecutorial districts, each served by one chief prosecutor ('District Attorney').

During the period we study, judges for both Superior and District Courts were elected in non-partisan elections.[47] For Superior Court judges, the length of term is 8 years. District Court judges serve 4-year terms. Prosecutors are also elected to 4-year terms in partisan elections. Judges and prosecutors are elected at the level of their respective districts. Some districts span multiple counties, and some fall *within* a county.

While most Superior Court judges are elected through the process described above, there are also a small number of Special Superior Court judges that are appointed by the governor. As of 2014, there were 12 Special Superior Court judges.

A.3 Texas

The data for Texas are derived from the Texas Computerized Criminal History System (CCH). The CCH is a statewide repository of criminal history data and includes data from various local criminal justice agencies, including arresting agencies, prosecuting agencies, and courts. Agencies are required to report data for all offenses that are Class B misdemeanors or greater. This includes all offenses that would potentially lead to a confinement sentence. The earliest records in the data date back to the early 20th century, though data quality and completeness improves over time. We focus on charges filed between 2000 and 2010.

The structure of the Texas data differs from the data collected from other states in that they are not derived solely from court records. In particular, the data only include court dispositions for offenses that are reported by some arresting agency. In our analysis, we drop offense records with no matched court data. We do this because we cannot code charge disposition in those cases. Of the arrests reported in the data over the years we study, about 85% of arrest records have matched court data. In the extract we obtained, merge rates fall after 2010. A 2011 audit from the Texas State Auditor's Office reports that courts may not submit records because: they encounter an error in the electronic submission process that is not reported back to the court; they lack the state identification numbers of arrest incident numbers required for merging; after an initial submission, they must correct or supply missing information manually via fax, resulting in lower submission rates.

To construct criminal histories for defendants, we use court data dating back to 1996. We stop at 1996 because the rate at which court records are matched to arrest records drops dramatically

[46] This rotation has occasionally been suspended due to budget constraints.
[47] The method of election was changed to partisan elections in 2017.

prior to 1996. Between 1996 and 2000, merge rates with court records are between 65-75%. Results are similar if we instead construct criminal histories using court data beginning in 1985 or 2000.

To measure charge outcomes, we take the original court disposition rather than any subsequent updates (for example, following a probation revocation).

We drop juvenile cases, and all cases for defendants below 16. We also drop cases where defendant demographic information, the offense, or court county are missing.

To match cases across individuals, we use the state identification number provided.

We exclude records from Loving County (population 67 in 2000) due to insufficient data. This leaves us with data from 253 counties.

In Texas, criminal cases are handled in District and County Courts. District Courts are courts of general jurisdiction, and handle all felony cases. There are 457 District Courts serving the 254 counties in the state. Each District Court corresponds to one judge. Most courts serve a single county. Some courts serve multiple, low-population counties. County Courts handle misdemeanors. There are 508 county courts.

Each county is served by at least one elected chief prosecutor ('County Attorney', 'District Attorney', or 'Criminal District Attorney'). In some counties, felony and misdemeanor cases are led by distinct chief prosecutors. Some prosecutors serve multiple counties.

Judges for both District and County Courts are elected in partisan elections. The length of term is 4 years. Prosecutors are also elected to 4 year terms in partisan elections.

A.4 Virginia

The data for Virginia are derived from administrative records from Virginia's Office of the Executive Secretary.

We restrict to felony and misdemeanor non-traffic offenses. We drop charges with missing dispositions, which appear to generally reflect charges that are on-going. We drop probation violations, and records that indicate intermediate outcomes, such as the transfer of a charge from one court to another. For misdemeanor charges that result in *de novo* appeals that send charges from a District Court to a Circuit Court, the data occasionally include both District Court and Circuit Court records, when only the Circuit Court record is relevant for sentencing. We drop District Court records for such appeals, matching based on defendant name, day and month of birth, and charge, and restricting to District Court records with guilty dispositions.

To match multiple cases to individuals, we group defendants based on full name and the day and month of birth. The Virginia data exclude year of birth.

The data do not include records from Alexandria or Fairfax. This leaves us with data from 118 cities and counties.

In Virginia, criminal cases are handled in Circuit and District Courts. Circuit Courts handle all felony cases. District Courts hear all criminal cases involving misdemeanors. There are Circuit and District Courts in every city and county. Circuit Courts are divided into 31 circuits. District Courts are divided into 32 districts. Each city and county is also served by one chief prosecutor ('Commonwealth's Attorney').

Circuit Court judges are appointed to 8-year terms by a majority of both houses of the General Assembly. District Court judges are also appointed by the legislature, but to 6-year terms. Prosecutors are elected to 4-year terms via partisan elections.

B Appendix: Criminal History Measurement

B.1 Alabama

In Alabama, the only legally required use of defendant criminal history as part of the sentencing decision is based on the Habitual Felony Offender Act. This legislation provides sentence enhancements as a function of the felony class charged and a defendant's number of prior felony convictions. Specifically, defendants are assigned to one of four categories and we use the same categories to define criminal history for misdemeanor defendants. While the State of Alabama also provides sentencing worksheets, these worksheets were voluntary during the 2000-2010 study period and so have not been used in the construction of defendant criminal histories.

B.2 North Carolina

In North Carolina, administrative court records explicitly record the number of points that a defendant has accrued prior to arrest as well as the class of each charged offense. The North Carolina Sentencing and Policy Advisory Commission publishes a matrix that presents presumptive sentence ranges as a function of offense class and the number of prior points for defendants charged with felonies and as a function of prior convictions and offense class for defendants charged with misdemeanors. Prior points categories are defined as follows: (1) 0-1 prior points, (2) 2-5 prior points, (3) 6-9 prior points, (4) 10-13 prior points, (5) 14-17 prior points, (6) 18+ prior points. Points are assigned based on the severity of past convictions and range from 10 points for a prior Class A felony conviction to 1 point for a prior Class 1 misdemeanor conviction. An additional point is assigned for offenders who have been previously convicted of a similar offense and for offenders who commit an offense while on supervised or unsupervised probation, parole, or while incarcerated. For misdemeanor offenses, the corresponding matrix provides presumptive sentences based on misdemeanor offense class and number of prior convictions (0, 1-4, 5+). Though these guidelines generate only three criminal history bins for defendants charged with misdemeanor offenses, we have verified that results are robust to employing a more continuous measure of criminal history based on number of past convictions, number of past incarceration sentences, and length of past incarceration sentences.

The presumptive sentence matrix for felonies was revised in 2009 and we rely on the updated matrix given that the North Carolina data included in the analysis sample are from 2007-2014. In practice, the revisions to the matrix were marginal; the prior point cutoffs associated with each criminal history bin were increased by a single point. While structured sentencing guidelines provide alternative ranges when aggravating or mitigating factors are present, judges otherwise have limited discretion to impose sentences outside of the official range.

B.3 Texas

In Texas, criminal histories are constructed as a function of offense class, number and severity of prior convictions, and, in a small number of instances, prior offense type. These criminal histories are defined to follow legislated conditions that trigger sentence enhancements. Under state law, a felony defendant with two prior felony convictions is subject to a mandatory minimum sentence of 25 years. First degree felony defendants are subject to a minimum sentence of 15 years if they have a previous (non-state jail) felony conviction and are subject to imprisonment for life if they are convicted of an aggravated sexual assault offense and have a previous conviction for a violent sexual offense. In addition, second and third degree felony defendants with a previous (non-state jail) felony conviction are subject to first and second degree felony punishments, respectively. State jail felony defendants are subject to third degree felony punishments if they have been previously convicted of two state jail felonies and to second degree felony punishments if they have been previously convicted of a non-state jail felony. Punishment for a third degree felony is also imposed for defendants with prior convictions for specifically-listed offenses. Class A misdemeanor defendants are subject to punishment enhancement if they have been previously convicted of a class A misdemeanor or a felony. Class B misdemeanor defendants are subject to punishment enhancement if they have been previously convicted of a class A misdemeanor, a class B misdemeanor, or a felony.

B.4 Virginia

In Virginia, judges maintain significant discretion in the sentencing of criminal defendants but rely on worksheets that calculate risk points based on charges faced and past convictions and incarceration sentences, among other factors. While the sentence ranges recommended based on these worksheets are voluntary, judges comply with the recommended ranges in 80% of non-jury cases. In total, there are 17 worksheets, corresponding to the most common offenses committed. For each worksheet, Section A is completed to determine whether an incarceration sentence is likely to be recommended and then Section B or C is completed based on the results of Section A to determine the specific punishment recommendation. We define criminal history based on the number of Section A risk points associated with a given charge. For charges that are not covered by worksheets, we apply the modal assignment of points based on past criminal history. In practice, the estimated number of risk points is measured with error since certain aggravating/mitigating factors cannot be determined based on the available data (juvenile convictions, victim age, weapon use, etc.). Sections B and C rely on many of the same factors that are included in Section A as well as a number of risk factors that cannot be measured using our data, and so we elect to define criminal history by offense type and estimated number of Section A risk points.

C Appendix: Mover Validation Exercise

C.1 Constructing Predictions in Mover Validation Exercise

Formally, we take first-differences of equation (1) to model the *change* in charge outcomes for a defendant moving from jurisdiction A to B:

$$
\begin{aligned}
\Delta_i Y_{ict} &= \Delta_i \tau_{cth(i,t)} + \Delta_i X_i \gamma^X + \Delta_i Z_{it} \gamma^Z + \Delta_i \theta_{j(i,c,t)} + \Delta_i \epsilon_{ict} \\
&= \Delta_i \tau_{cth(i,t)} + \Delta_i Z_{it} \gamma^Z + \Delta_i \theta_{j(i,c,t)} + \Delta_i \epsilon_{ict} \\
\Delta_i Y_{ict} - \Delta_i \tau_{cth(i,t)} - \Delta_i Z_{it} \gamma^Z &= \Delta \theta_{j(i,c,t)} + \Delta_i \epsilon_{ict}
\end{aligned}
$$

Appendix Figure C1 plots the distribution of $\Delta\theta_{j(i,c,t)}$ for movers, separately by state. For each defendant i we plug in coefficients for $\tau_{cth(i,t)}$ and γ^Z as well as punishment severity θ_j from the model estimated using the 9 subsets that do not include defendant i and estimate the following model for movers:

$$
\underbrace{\Delta Y_{ict} - \Delta \widehat{\tau_{cth(i,t)}} - \Delta Z_{it} \widehat{\gamma^Z}}_{\text{adjusted change in confinement}} = \alpha + \beta \underbrace{\Delta \widehat{\theta_{j(i,t)}}}_{\text{change in punishment severity}} + \xi_{ict}
$$

adding a constant term α to allow for systematic prediction error. A β coefficient of one indicates that the punishment severity estimates are unbiased.

C.2 Do Mover Defendants Sort on Match Effects?

We test whether mover defendants appear to sort on two types of match effects: jurisdiction by crime type and jurisdiction by criminal history interactions. In particular, in two separate exercises we test whether mover defendants that move to jurisdictions with larger estimated punishment severity also (a) commit offenses or (b) have criminal histories that are punished particularly harshly or leniently in that jurisdiction. To do this, we first estimate

$$
y_{ict} = \tau_{cth(i,t)} + x_i \gamma^x + z_{it} \gamma^z + \theta^M_{j(i,c,t),k(i,c,t)} + \epsilon_{ict} \tag{C.1}
$$

using the same split-sample procedure described in Section 3.2, where $\theta^M_{j(i,c,t),k(i,c,t)}$ are (a) crime type match effects or (b) criminal history match effects, or fixed effects for each (a) jurisdiction by crime type interaction or (b) jurisdiction by criminal history interaction. We use four crime types: violent, property, drug, and other. We use three categories of criminal history. For each state, using state-specific criminal history scores, we calculate the median criminal history among those with any criminal history. We then divide defendants into three groups: those with zero criminal history, those with criminal history below the conditional median, and those with criminal history above the conditional median. We then take the sample of mover defendant charges used to construct Figure 2 Panel A and plot the change in (a) crime type match effects or (b) criminal history match effects against the change in estimated punishment severity. In the absence of sorting, changes in punishment severity should predict changes in match effects without bias.

The results are depicted in Panel A (crime type) and Panel B (criminal history) of Figure C2. In both cases, we find little to no evidence of sorting based on match effects. All points fall on or very near the $45°$ line. For crime type, the slope coefficient estimate is 1.00. For criminal history, the slope coefficient estimate is 1.04.

C.3 Across-Jurisdiction Variance Decomposition

We present a second decomposition of across-jurisdiction variation in confinement rates. Motivated by the fact that

$$Var(\bar{y}_j) = Var(\theta_j) + Var(\bar{\gamma}_j) + Var(\bar{\tau}_j) + 2Cov(\theta_j, \bar{\gamma}_j) + 2Cov(\theta_j, \bar{\tau}_j) + 2Cov(\bar{\gamma}_j, \bar{\tau}_j), \quad \text{(C.2)}$$

we compute and report the sample analog for each term. In estimating each variance and covariance term, we follow Finkelstein et al. (2016) and use a split-sample approach to correct for sampling error. We randomly assign defendants into two subsamples of approximately equal size and estimate equation (2) separately using each subsample. We estimate the variance of $\hat{\theta}_j$ using the covariance between the $\hat{\theta}_j$ estimates derived from the two subsamples. We take an analogous approach to estimate the variances of $\hat{\gamma}_j$ and $\hat{\tau}_j$. We compute the covariance between $\hat{\theta}_j$ and $\hat{\gamma}_j$ as the average of the covariances between $\hat{\theta}_j$ from one subsample and $\hat{\gamma}_j$ from the other subsample. We compute all other covariance terms analogously. We then compute $Var(\bar{y}_j)$ based on our estimated variance and covariance terms following equation (C.2). We compute the correlation between $\hat{\theta}_j$ and $\hat{\gamma}_j$ using our estimated variances of $\hat{\theta}_j$ and $\hat{\gamma}_j$ and covariance between $\hat{\theta}_j$ and $\hat{\gamma}_j$.

Using this decomposition, we also ask what share of cross-jurisdiction variation in confinement rates would be eliminated in a counterfactual where jurisdiction effects θ_j were equalized across jurisdictions. This share corresponds to

$$1 - \frac{Var(\bar{\gamma}_j) + Var(\bar{\tau}_j) + 2Cov(\bar{\gamma}_j, \bar{\tau}_j)}{Var(\bar{y}_j)}$$

because when jurisdiction effects are equalized

$$Var(\theta_j) = Cov(\theta_j, \bar{\gamma}_j) = Cov(\theta_j, \theta_j) = 0.$$

Similarly, we ask what share of cross-jurisdiction variation in confinement rates would be eliminated if defendant effects or charge effects were equalized. Note that these terms can be negative, in which case equalizing a component across jurisdictions would *increase* cross-jurisdiction variation in confinement rates.

Table C1 reports the results for this decomposition. We find that 64%-93% of variance would be eliminated if jurisdiction effects were equalized. By contrast, variance would be reduced by a small amount or even increase if defendant or charge effects were equalized. In Alabama, North Carolina, and Virginia, jurisdiction and defendant effects are negatively correlated, with the correlation ranging from -0.355 to -0.272. In Texas, jurisdiction and defendant effects are essentially uncorrelated.

Finally, we compute the share of defendant effects that are explained by defendant observables, which include race, sex, and age. Observables explain 2.0%, 3.1%, 4.4%, and 1.3% of defendant effects in Alabama, North Carolina, Texas, and Virginia, respectively. Observable defendant characteristics explain little of the variation in defendant effects in part because we observe few charges per defendant and the outcome is an indicator variable.

D Appendix: Additional Data Sources

D.1 Statewide Ballot Measures

As we discuss in Section 4.2.1, we construct a proxy for local voter punishment preferences using jurisdiction-level data on support for statewide ballot measures related to the punishment of criminals and the rights of the accused. Data on the universe of potentially relevant ballot measures were generously shared with us by Claire Lim, James Snyder, Jr., and David Strömberg. To construct the measures used in our analysis (presented in Appendix Table A14), we limited the sample of ballot measures to exclude those related to victims' rights as these measures were not well-suited for capturing local attitudes towards punishment. In addition, the data contained two closely-related ballot measures from 2005 and 2007 in Texas. We dropped voting data for the second measure (Proposition 13 (2007)) since voter turnout for the 2007 measure was especially low (only 50% as high as 2005 turnout). We were then left with one ballot measure for each of the four states included in our sample: (1) Amendment 3 (1996) from Alabama, which removed the prohibition on guilty pleas within 15 days of arrest in non-capital felony cases; (2) Amendment 2 (1996) from North Carolina, which expanded the types of punishment that could be imposed on convicted criminals; (3) Proposition 4 (2005) from Texas, which authorized the denial of bail to a criminal defendant who violates a condition of the defendant's release pending trial; (4) Proposition 3 (1996) from Virginia, which authorized the legislature to allow the state the right of an appeal in all cases, including criminal cases. The vote share measure employed in our analysis is the county-level share of voters who supported each ballot measure.

D.2 State Court Processing Statistics Data

As we discuss in Section 4.2.1, we use data from the State Court Processing Statistics Data series to assess the generalizability of our findings. These data include felony cases filed in the nation's most populous counties in even numbered years from 1990-2006 and 2009. The data include a subset of felony cases filed in May of the referenced year in each county. Importantly, the data include cases that were ultimately dismissed or did not otherwise result in conviction. To construct county-specific punishment severity measures, we pooled all included data and then estimated the case-level equivalent of equation (1). Specifically, we regressed an indicator for any confinement sentence on a set of demographic controls (defendant race, gender, age and age squared), year-by-offense category-by-criminal history fixed effects, and jurisdiction fixed effects. Offense category was defined by the most serious arrest charge in combination with the number of total charges included in the case and criminal history was defined based on the total number of prior convictions. Case-level weights were applied to account for within-jurisdiction sampling from the universe of felony cases filed in each jurisdiction in May of the referenced year. As in our benchmark analysis, the estimated jurisdiction fixed effects were adjusted to reflect the predicted confinement rate for each jurisdiction based on the overall composition of included cases. The logged state-level mean punishment severity included in a subset of specifications in Appendix Table A15 are simply the logged values of the (unweighted) state-level averages of these adjusted jurisdiction-level measures.

Table A1: Charges per Reported Crime Across Jurisdictions

	Alabama	North Carolina	Texas	Virginia
Charges per Crime UCR Part I:				
Mean	0.245	0.664	0.203	0.475
SD	(0.204)	(0.517)	(0.179)	(0.193)
N Jurisdictions	67	98	253	118

Notes: 'Charges per Crime UCR Part I' is the total number of recorded charges for UCR Part I offenses in a county and year divided by total reported UCR Part I offenses in that county and year, averaged across years.

Table A2: Coefficient Estimates from Punishment Severity Models, Case-Level

Outcome: Confinement	Alabama	North Carolina	Texas	Virginia
Black	0.043**	0.026**	0.074**	0.035**
	(0.001)	(0.000)	(0.001)	(0.001)
Hispanic		0.043**	0.058**	
		(0.001)	(0.001)	
Male	0.053**	0.035**	0.099**	0.046**
	(0.001)	(0.000)	(0.000)	(0.001)
Age	0.001**	0.006**	0.014**	
	(0.000)	(0.000)	(0.000)	
Age^2 $\times$ 100	-0.002**	-0.006**	-0.016**	
	(0.000)	(0.000)	(0.000)	
Criminal History $\times$ Charge $\times$ Year Fixed Effects	✓	✓	✓	✓
Jurisdiction Fixed Effects	✓	✓	✓	✓
N Cases	1,221,317	3,984,894	4,931,314	1,777,549
Adjusted R^2	0.220	0.150	0.217	0.240
Mean Confinement	0.248	0.112	0.403	0.226

Notes: Table presents coefficients from state-specific estimates of equation (1) estimated at the case-level. More details on how case-level and single charge case estimates are produced are discussed in Section 3.1.3. Missing values reflect characteristics that are unavailable for particular states.
Standard errors clustered by defendant in parentheses. ˜ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A3: Coefficient Estimates from Punishment Severity Models, Single Charge Cases

Outcome: Confinement	Alabama	North Carolina	Texas	Virginia
Black	0.041**	0.021**	0.072**	0.035**
	(0.001)	(0.000)	(0.001)	(0.001)
Hispanic		0.039**	0.056**	
		(0.001)	(0.001)	
Male	0.046**	0.031**	0.096**	0.041**
	(0.001)	(0.000)	(0.001)	(0.001)
Age	0.001**	0.005**	0.014**	
	(0.000)	(0.000)	(0.000)	
Age^2 $\times$ 100	-0.002**	-0.005**	-0.016**	
	(0.000)	(0.000)	(0.000)	
Criminal History $\times$ Charge $\times$ Year Fixed Effects	✓	✓	✓	✓
Jurisdiction Fixed Effects	✓	✓	✓	✓
N Cases	889,433	2,983,923	4,241,432	1,333,691
Adjusted R^2	0.206	0.103	0.208	0.203
Mean Confinement	0.210	0.089	0.377	0.187

Notes: Table presents coefficients from state-specific estimates of equation (1) restricted to single charge cases. More details on how case-level and single charge case estimates are produced are discussed in Section 3.1.3. Missing values reflect characteristics that are unavailable for particular states.

Standard errors clustered by defendant in parentheses. ~ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A4: Comparing Punishment Severity Estimates: Charge-Level, Case-Level, and Single Charge Cases

	Alabama	North Carolina	Texas	Virginia
Case-Level:				
Avg. Confinement Rate (%)	22.3	10.7	23.2	23.4
SD of Punishment Severity	11.4	2.8	11.3	4.8
Single Charge Cases:				
Avg. Confinement Rate (%)	18.2	8.3	20.8	18.1
SD of Punishment Severity	10.9	2.4	10.8	4.5
Correlations:				
Baseline vs. Case-Level	0.963	0.985	0.993	0.929
Baseline vs. Single Charge	0.945	0.945	0.986	0.891
Case-Level vs. Single Charge	0.992	0.977	0.995	0.986
Number of Jurisdictions	67	100	253	118

Notes: Details on how case-level and single charge case estimates are produced are discussed in Section 3.1.3.

Table A5: Descriptive Statistics: Stayers versus Movers

State:	Alabama			North Carolina			Texas			Virginia		
	All	Multiple Cases		All	Multiple Cases		All	Multiple Cases		All	Multiple Cases	
		Stayer	Mover		Stayer	Mover		Stayer	Mover		Stayer	Mover
Male	70.2	69.9	74.8	73.7	77.3	78.6	77.2	81.6	81.7	70.9	73.6	76.6
Black	39.7	42.9	26.0	44.5	49.3	37.3	25.1	30.8	22.8	43.3	46.1	42.8
Hispanic				5.2	3.5	1.8	32.9	35.0	25.4			
Age	33.1	32.4	32.1	32.1	31.9	30.2	31.3	30.4	30.4			
	(11.2)	(10.5)	(9.4)	(12.4)	(12.0)	(10.4)	(11.3)	(10.5)	(9.7)			
Felony	35.2	36.0	36.6	25.6	32.3	32.2	30.8	34.0	31.9	35.4	46.7	51.2
Property	16.1	17.2	21.0	27.6	30.4	38.8	20.9	20.6	24.7	30.1	37.6	42.7
Violent	10.8	8.8	6.7	14.8	14.5	11.0	13.2	13.6	10.0	11.4	10.7	10.2
Drug	17.7	15.9	15.8	19.7	21.7	19.2	21.7	23.8	22.1	14.7	16.7	15.6
Other	55.4	58.2	56.6	38.0	33.5	31.0	44.2	42.0	43.1	43.8	35.1	31.5
Dropped	41.5	37.9	35.9	61.9	63.0	59.1	23.0	20.2	20.7	44.0	44.4	41.8
Convicted	56.3	60.6	62.7	35.7	34.9	39.2	52.8	60.6	60.8	50.8	51.3	53.9
Probation	27.1	31.2	30.8	15.6	14.2	15.4	32.8	25.9	27.2	11.2	12.2	11.9
Confinement	20.1	25.5	25.9	7.7	8.8	10.1	37.8	48.3	45.1	16.5	20.1	25.0
Sentence $\geq$ 90 Days	15.2	19.1	19.8	3.1	3.9	4.5	26.6	31.7	31.3	7.9	10.8	13.7
N Defendants	676,253	146,094	51,166	1,628,326	491,639	211,925	2,161,639	608,069	427,002	991,886	175,206	117,025
N Charges	1,504,992	739,123	349,216	4,012,569	2,580,369	1,729,714	3,947,711	2,201,407	1,928,737	1,891,035	793,092	722,262
N Cases	1,010,555	480,396	210,762	2,882,130	1,745,443	1,102,764	3,373,485	1,819,915	1,557,829	1,336,562	519,882	440,987
Charges per Defendant	2.2	5.1	6.8	2.5	5.2	8.2	1.8	3.6	4.5	1.9	4.5	6.2
	(3.6)	(6.0)	(8.5)	(4.1)	(6.4)	(9.8)	(1.8)	(2.6)	(3.3)	(3.3)	(6.5)	(6.8)
Cases per Defendant	1.5	3.3	4.1	1.8	3.6	5.2	1.6	3.0	3.6	1.3	3.0	3.8
	(1.7)	(2.9)	(4.0)	(2.1)	(3.2)	(4.6)	(1.3)	(1.8)	(2.3)	(1.2)	(2.4)	(2.8)
Charges per Case	1.5	1.5	1.7	1.4	1.5	1.6	1.2	1.2	1.2	1.4	1.5	1.6
	(1.8)	(1.8)	(2.0)	(1.5)	(1.7)	(2.0)	(0.6)	(0.6)	(0.7)	(2.0)	(2.5)	(1.9)

Notes: Standard deviation in parentheses. Missing values reflect characteristics that are unavailable for particular states. ‘Other’ offenses include crimes against society and offenses we are unable to classify due to miscoding.

Table A6: Pre- and Post-Move Charges

State:	Alabama		North Carolina		Texas		Virginia	
	Stayers	Movers	Stayers	Movers	Stayers	Movers	Stayers	Movers
Neighboring Counties (%)		68.7		63.5		53.5		65.2
Same Offense Type (%)	69.3	50.4	54.2	38.0	40.9	37.5	55.8	41.6
Pre-Move Charge (%):								
Property	15.0	19.3	33.2	32.1	22.5	26.3	39.6	37.5
Violent	8.9	9.1	15.4	12.0	13.2	9.6	9.7	10.3
Drug	12.6	17.4	18.4	17.9	23.5	22.1	17.8	16.4
Other	63.6	54.2	33.1	38.0	40.8	42.1	32.9	35.8
Post-Move Charge (%):								
Property	14.9	20.2	33.2	32.1	21.7	25.1	37.1	36.7
Violent	8.5	9.3	15.1	12.4	13.8	10.1	10.0	11.1
Drug	12.6	17.6	18.7	18.0	23.0	21.2	17.6	16.3
Other	64.0	53.0	33.0	37.5	41.5	43.6	35.3	35.8
N Case Pairs	416,353	75,640	2,187,477	328,134	1,656,521	679,075	589,869	72,649

Notes: This table describes charges for mover and stayer defendants as described in Section 3.2.

Table A7: Descriptive Statistics for County Characteristics

	Alabama	North Carolina	Texas	Virginia	Observations
Black Population Share	0.283	0.216	0.069	0.201	538
	(0.222)	(0.168)	(0.073)	(0.172)	
Black Defendant Share	0.391	0.347	0.140	0.322	538
	(0.238)	(0.224)	(0.125)	(0.216)	
Log Pop. Density	3.996	4.662	3.013	4.916	538
	(0.890)	(0.913)	(1.656)	(1.595)	
Log Average HH Income	9.656	9.775	9.673	9.851	538
	(0.168)	(0.161)	(0.202)	(0.196)	
Gini Index	0.446	0.429	0.431	0.392	471
	(0.071)	(0.086)	(0.073)	(0.079)	
Fraction Males Aged 15-29	0.101	0.101	0.103	0.098	538
	(0.014)	(0.023)	(0.026)	(0.031)	
Violent Crime Rate,	0.500	0.171	-0.107	-0.192	531
2000 (Standardized)	(1.761)	(0.875)	(0.739)	(0.918)	
Violent Crime Rate Growth,	-0.345	-0.356	0.153	0.171	530
1970-1990 (Standardized)	(1.403)	(1.151)	(0.842)	(0.756)	

Notes: Excluding violent crime rate growth from 1970 to 1990, characteristics are measured in 2000.

Table A8: Punishment Severity and Charges Recorded

Outcome:	Charge to Crime Ratio			Log Relative Punishment Severity			Log Relative Punishment Severity, Violent and Property Crime Charges	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Charge to Crime Ratio				0.015	0.016	0.000	0.017	0.001
(Standardized)				(0.016)	(0.016)	(0.016)	(0.019)	(0.018)
Black Share of Population	-2.243**	-0.253	0.012		0.914**	0.854**		0.856**
	(0.836)	(0.782)	(0.717)		(0.261)	(0.280)		(0.309)
Black Share of Population, Squared	3.278*	0.845	-0.273		-1.263**	-1.464**		-1.637**
	(1.275)	(1.087)	(1.004)		(0.402)	(0.451)		(0.474)
Log Population Density		-0.071	x	0.101**	0.099**	x	0.103**	x
		(0.043)		(0.021)	(0.021)		(0.027)	
Log Average Household Income		-0.580*	x	0.378*	0.361*	x	0.384˜	x
		(0.257)		(0.164)	(0.167)		(0.204)	
Gini Coefficient		-0.598	x	-0.040	-0.073	x	-0.071	x
		(0.573)		(0.233)	(0.227)		(0.254)	
Fraction Males Aged 15-29		-0.674	x	0.253	0.073	x	0.670	x
		(2.015)		(0.742)	(0.729)		(0.739)	
Violent Crime Rate, 2000		-0.217**	x	-0.027	-0.038*	x	-0.031	x
		(0.056)		(0.017)	(0.018)		(0.019)	
State FEs	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.008	0.113	0.159	0.190	0.202	0.233	0.164	0.189
Observations	535	535	535	535	535	535	535	535

Notes: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls. We standardize the charge to crime ratio to have mean zero and standard deviation one within states. Charge to crime ratio data are missing for three counties. In columns (7) and (8) the outcome is log relative punishment severity, but estimated only using violent and property crime charges.

'x' denotes inclusion of the covariate interacted with state fixed effects.

Robust standard errors in parentheses. ˜ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A9: Punishment Severity and Population Slave Share in 1860

Outcome:	Log Relative Punishment Severity				
	(1)	(2)	(3)	(4)	(5)
Slave Share	1.436**	0.864*	0.842*	0.716*	0.838*
	(0.430)	(0.354)	(0.356)	(0.359)	(0.410)
Slave Share,	-1.966**	-1.023*	-0.973*	-0.815˜	-1.095˜
Squared	(0.581)	(0.471)	(0.474)	(0.470)	(0.558)
Log Population Density		0.117**	0.125**	†	x
		(0.036)	(0.039)		
Log Average Household Income		0.242	0.236	0.272	x
		(0.228)	(0.238)	(0.246)	
Gini Coefficient		0.009	0.032	0.107	x
		(0.251)	(0.252)	(0.255)	
Fraction Males Aged 15-29		0.281	0.182	0.245	x
		(1.128)	(1.132)	(1.150)	
Violent Crime Rate Growth, 1970-1990		-0.031˜		-0.037˜	
		(0.019)		(0.019)	
Violent Crime Rate, 2000			-0.010	-0.006	x
			(0.020)	(0.020)	
State FEs	✓	✓	✓	✓	✓
Adjusted R^2	0.050	0.193	0.187	0.201	0.236
Observations	361	361	361	361	361

Notes: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls.
'†' denotes inclusion of a five-piece linear spline in log population density. 'x' denotes inclusion of the covariate interacted with state fixed effects.
Robust standard errors in parentheses. ˜ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A10: Race-Based Confinement Gaps

Outcome:	Black-White Log Relative Punishment Severity					
	(1)	(2)	(3)	(4)	(5)	(6)
Black Share of Population	0.429~	0.418*	0.254			
	(0.242)	(0.191)	(0.184)			
Black Share of Population, Squared	-0.741*	-0.796**	-0.375			
	(0.338)	(0.265)	(0.245)			
Black Share of Defendants				0.136	0.174	-0.004
				(0.257)	(0.206)	(0.197)
Black Share of Defendants, Squared				-0.184	0.269	0.059
				(0.299)	(0.257)	(0.248)
Log Population Density		0.005	x		0.007	x
		(0.015)			(0.014)	
Log Average Household Income		0.148	x		0.138	x
		(0.092)			(0.092)	
Gini Coefficient		0.129	x		0.104	x
		(0.144)			(0.143)	
Fraction Males Aged 15-29		0.529	x		0.607	x
		(0.399)			(0.392)	
Violent Crime Rate, 2000		0.001	x		0.002	x
		(0.009)			(0.010)	
State FEs	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.056	0.055	0.050	0.049	0.049	0.048
Observations	533	533	533	533	533	533

Notes: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls. 'x' denotes inclusion of the covariate interacted with state fixed effects.
Robust standard errors in parentheses. ~ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A11: Punishment Severity and Racial Heterogeneity, Within-Defendant and Subgroup Estimates

Outcome:	Within-Defendant (1)	Black (2)	White (3)	First Offense (4)	Subsequent Offense (5)	Violent Crime (6)	Property Crime (7)	Drug Crime (8)
Black Share of Population	0.862**	1.193**	0.802**	1.115**	0.891**	1.103**	0.930**	0.843**
	(0.316)	(0.280)	(0.263)	(0.287)	(0.270)	(0.240)	(0.266)	(0.312)
Black Share of Population, Squared	-1.043*	-1.609**	-0.857*	-1.518**	-1.242**	-1.731**	-1.548**	-1.532**
	(0.514)	(0.407)	(0.401)	(0.434)	(0.444)	(0.403)	(0.410)	(0.532)
Log Population Density	0.126**	0.113**	0.110**	0.118**	0.105**	0.100**	0.122**	0.104**
	(0.024)	(0.020)	(0.019)	(0.021)	(0.021)	(0.020)	(0.026)	(0.024)
Log Average Household Income	0.234	0.168	0.280*	0.349*	0.315~	0.134	0.271	0.114
	(0.158)	(0.123)	(0.123)	(0.157)	(0.164)	(0.111)	(0.176)	(0.173)
Gini Coefficient	-0.033	-0.265	-0.105	-0.042	-0.006	-0.163	-0.233	-0.033
	(0.287)	(0.220)	(0.224)	(0.257)	(0.221)	(0.214)	(0.247)	(0.243)
Fraction Males Aged 15-29	-0.094	0.659	0.088	-0.057	0.062	-0.101	-0.128	-0.352
	(0.978)	(0.659)	(0.624)	(0.686)	(0.730)	(0.663)	(0.873)	(0.797)
Violent Crime Rate, 2000	-0.081**	-0.039*	-0.042*	-0.062**	-0.042*	-0.052**	-0.021	-0.027
	(0.025)	(0.017)	(0.017)	(0.020)	(0.018)	(0.018)	(0.020)	(0.020)
State FEs	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.139	0.189	0.213	0.188	0.193	0.177	0.205	0.163
Observations	536	533	538	532	538	535	536	537

Notes: The outcome measure in each column is the log relative punishment severity, derived by estimating equation (2) (for column (1)) and derived by estimating equation (1) for specific subgroups (in subsequent columns). Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls. There are fewer than 538 observations for some outcomes because there are jurisdictions where the predicted confinement rate is either missing (due to insufficient data) or below zero (and so the log transformation is undefined) for the relevant subset of charges or defendants.

Robust standard errors in parentheses. ~ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A12: Punishment Severity and Racial Heterogeneity, Weighted by Jurisdiction Population

Outcome:	Log Relative Punishment Severity									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black Share of Population	2.751**	1.341**	1.452**	1.439**	1.505**					
	(0.664)	(0.428)	(0.441)	(0.391)	(0.407)					
Black Share of Population, Squared	-4.585**	-2.307**	-2.309**	-2.124**	-2.614**					
	(1.080)	(0.733)	(0.755)	(0.705)	(0.681)					
Black Share of Defendants						2.791**	1.959**	1.980**	1.812**	1.714**
						(0.477)	(0.403)	(0.411)	(0.375)	(0.363)
Black Share of Defendants, Squared						-3.165**	-2.517**	-2.467**	-2.148**	-2.033**
						(0.566)	(0.537)	(0.560)	(0.544)	(0.514)
Log Population Density		0.069**	0.081**	†	x		0.068**	0.076**	†	x
		(0.025)	(0.027)				(0.023)	(0.025)		
Log Average Household Income		0.308*	0.288*	0.330*	x		0.156	0.137	0.180	x
		(0.135)	(0.139)	(0.131)			(0.128)	(0.134)	(0.130)	
Gini Coefficient		0.048	0.183	0.276	x		0.242	0.329	0.396	x
		(0.318)	(0.310)	(0.300)			(0.306)	(0.300)	(0.286)	
Fraction Males Aged 15-29		-0.626	-0.712	-0.938	x		-0.743	-0.824	-1.036	x
		(0.766)	(0.771)	(0.747)			(0.698)	(0.702)	(0.672)	
Violent Crime Rate Growth, 1970-1990		-0.014		-0.028			-0.014		-0.025	
		(0.028)		(0.026)			(0.026)		(0.025)	
Violent Crime Rate, 2000			-0.033	-0.029	x			-0.023	-0.021	x
			(0.023)	(0.022)				(0.024)	(0.024)	
State FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.385	0.492	0.494	0.529	0.556	0.460	0.530	0.530	0.555	0.571
Observations	538	538	538	538	538	538	538	538	538	538

Notes: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. Observations weighted by jurisdiction population in 2000. For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls.
'†' denotes inclusion of a five-piece linear spline in log population density. 'x' denotes inclusion of the covariate interacted with state fixed effects.
Robust standard errors in parentheses. ~ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A13: Punishment Severity and Racial Heterogeneity, Degree of Selection on Unobservables

Outcome:	Log Relative Punishment Severity					
	(1)	(2)	(3)	(4)	(5)	(6)
Black Share of Population	1.644**	0.716**	0.895**			
	(0.297)	(0.268)	(0.279)			
δ degree of selection, relative to Column (1)			2.1			
δ degree of selection, relative to Column (2)			-5.6			
Black Share of Population, Squared	-2.763**	-1.287**	-1.494**			
	(0.501)	(0.441)	(0.450)			
δ degree of selection, relative to Column (1)			2.1			
δ degree of selection, relative to Column (2)			-12.4			
Black Share of Defendants				2.147**	1.314**	1.424**
				(0.270)	(0.271)	(0.279)
δ degree of selection, relative to Column (4)						2.1
δ degree of selection, relative to Column (5)						11.0
Black Share of Defendants, Squared				-2.705**	-1.745**	-1.905**
				(0.353)	(0.355)	(0.389)
δ degree of selection, relative to Column (4)						2.1
δ degree of selection, relative to Column (5)						3.8
Log Population Density		x	x		x	x
Log Average Household Income			x			x
Gini Coefficient			x			x
Fraction Males Aged 15-29			x			x
Violent Crime Rate, 2000			x			x
State FEs	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.049	0.211	0.238	0.111	0.240	0.266
Observations	538	538	538	538	538	538

Notes: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. For covariates that are missing for some jurisdictions (crime rates and Gini index), we set missing values to zero and include indicators for missing data for each of these covariates as additional controls. δ degree of selection values are constructed as in Oster (2019) and reflect the degree of selection on unobservables (relative to observables) needed to explain estimates. We follow the examples provided in Oster (2019) and assume $R_{max} = 1.3$ times the R^2 from the saturated model.
'x' denotes inclusion of the covariate interacted with state fixed effects.
Robust standard errors in parentheses. ˜ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A14: Ballot Measure Voting and Racial Heterogeneity

Outcome:	Log Relative Punishment Severity	Harshness Vote Share (Standardized)								Log Relative Punishment Severity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Harshness Vote Share	0.053**									
	(0.020)									
Black Share of Population		2.932**	1.975*	2.064**	1.772*					
		(0.851)	(0.776)	(0.787)	(0.823)					
Black Share of Population, Squared		-3.902*	-2.338~	-2.222	-1.719					
		(1.591)	(1.376)	(1.387)	(1.453)					
Black Share of Defendants						2.641**	1.387*	1.349~	1.346~	
						(0.708)	(0.680)	(0.704)	(0.752)	
Black Share of Defendants, Squared						-2.591*	-1.257	-1.036	-1.216	
						(1.029)	(0.979)	(1.047)	(1.132)	
Log Population Density			0.044	†	x		0.036	†	x	
			(0.058)				(0.059)			
Log Average Household Income			0.773*	0.732~	x		0.691~	0.653~	x	
			(0.382)	(0.381)			(0.374)	(0.375)		
Gini Coefficient			-2.166**	-2.125**	x		-2.059**	-2.024**	x	
			(0.748)	(0.765)			(0.765)	(0.778)		
Fraction Males Aged 15-29			-2.186	-2.212	x		-2.103	-2.101	x	
			(2.187)	(2.257)			(2.157)	(2.229)		
Violent Crime Rate Growth, 1970-1990			0.070	0.064			0.069	0.064		
			(0.050)	(0.051)			(0.051)	(0.052)		
Violent Crime Rate, 2000			0.166**	0.167**	x		0.169**	0.170**	x	
			(0.054)	(0.054)			(0.054)	(0.054)		
Republican Vote Share, 2000										0.730**
										(0.224)
State FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.009	0.021	0.110	0.114	0.110	0.033	0.108	0.111	0.107	0.035
Observations	538	538	538	538	538	538	538	538	538	538

Notes: Harshness Vote Share is constructed using jurisdiction-level data on support for statewide ballot measures related to the punishment of criminals and the rights of the accused. Additional details regarding the data source are provided in Appendix D. Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio.
'†' denotes inclusion of a five-piece linear spline in log population density. 'x' denotes inclusion of covariate interacted with state fixed effects.
Notes: Robust standard errors in parentheses. ~ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table A15: State Court Processing Statistics Analyses

Outcome:	Log State-level Mean Punishment Severity	Log County-level Punishment Severity		Log State-level Mean Punishment Severity	Log County-level Punishment Severity	
	(1)	(2)	(3)	(4)	(5)	(6)
Black Share of State Population	7.152~	10.146**	9.751**	22.879**	17.772*	16.017*
P-value	*0.052*	*0.001*	*0.006*	*0.000*	*0.036*	*0.040*
Black Share of State Population, Squared	-21.362*	-28.627**	-26.624*	-63.504**	-51.236*	-44.265*
P-value	*0.035*	*0.004*	*0.013*	*0.004*	*0.026*	*0.034*
Black Share of County Population		-0.579	-0.997		1.471	2.973
P-value		*0.492*	*0.241*		*0.492*	*0.356*
Black Share of County Population, Squared		0.164	1.110		-1.980	-3.886
P-value		*0.919*	*0.458*		*0.550*	*0.442*
Unit of Observation	State	County	County	State	County	County
Region FEs	✓	✓	✓			
Excludes 4 Sample States			✓			✓
Southern Region Only				✓	✓	✓
R^2	0.560	0.567	0.621	0.738	0.322	0.472
Observations	25	70	62	9	22	14

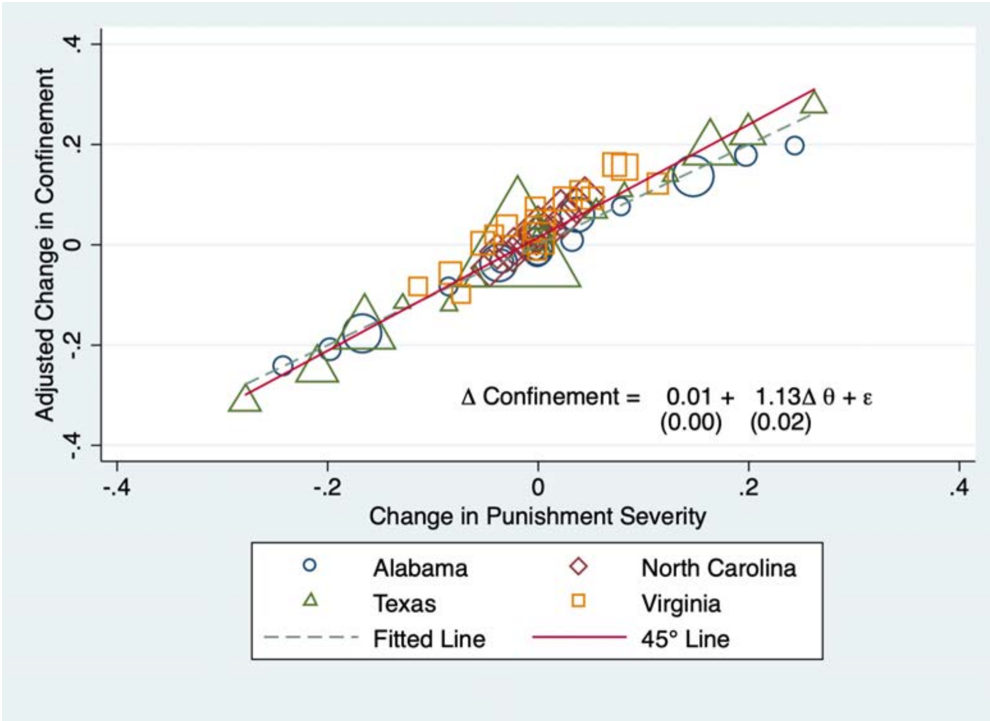
Notes: Punishment severity is constructed from case-level State Court Processing Statistics data on felony defendants. Additional details regarding the data source are provided in Appendix D.
P-values constructed by randomization inference in Columns 4-6.
~ significant at 10 percent level; * significant at 5 percent level; ** significant at 1 percent level.

Table C1: Across-Jurisdiction Variance Decomposition of Confinement Rates

	Alabama	North Carolina	Texas	Virginia
Across-jurisdiction variance of average:				
Confinement rates	123.68	5.49	140.88	30.37
Jurisdiction effects	121.57	6.03	128.78	31.90
Defendant effects	45.26	1.89	15.58	8.89
Charge effects	1.45	0.25	2.70	1.80
Across-jurisdiction covariance of average:				
Jurisdiction and defendant effects	-21.32	-0.92	1.85	-5.98
Jurisdiction and charge effects	0.22	-0.19	-0.70	0.76
Defendant and charge effects	-1.20	-0.23	-4.24	-0.89
Correlation of jurisdiction and defendant effects	-0.287	-0.272	0.041	-0.355
Share variance would be reduced if:				
Jurisdiction effects were made equal	0.64	0.69	0.93	0.71
Defendant effects were made equal	0.00	-0.06	0.08	-0.16
Charge effects were made equal	0.00	-0.11	-0.05	0.05

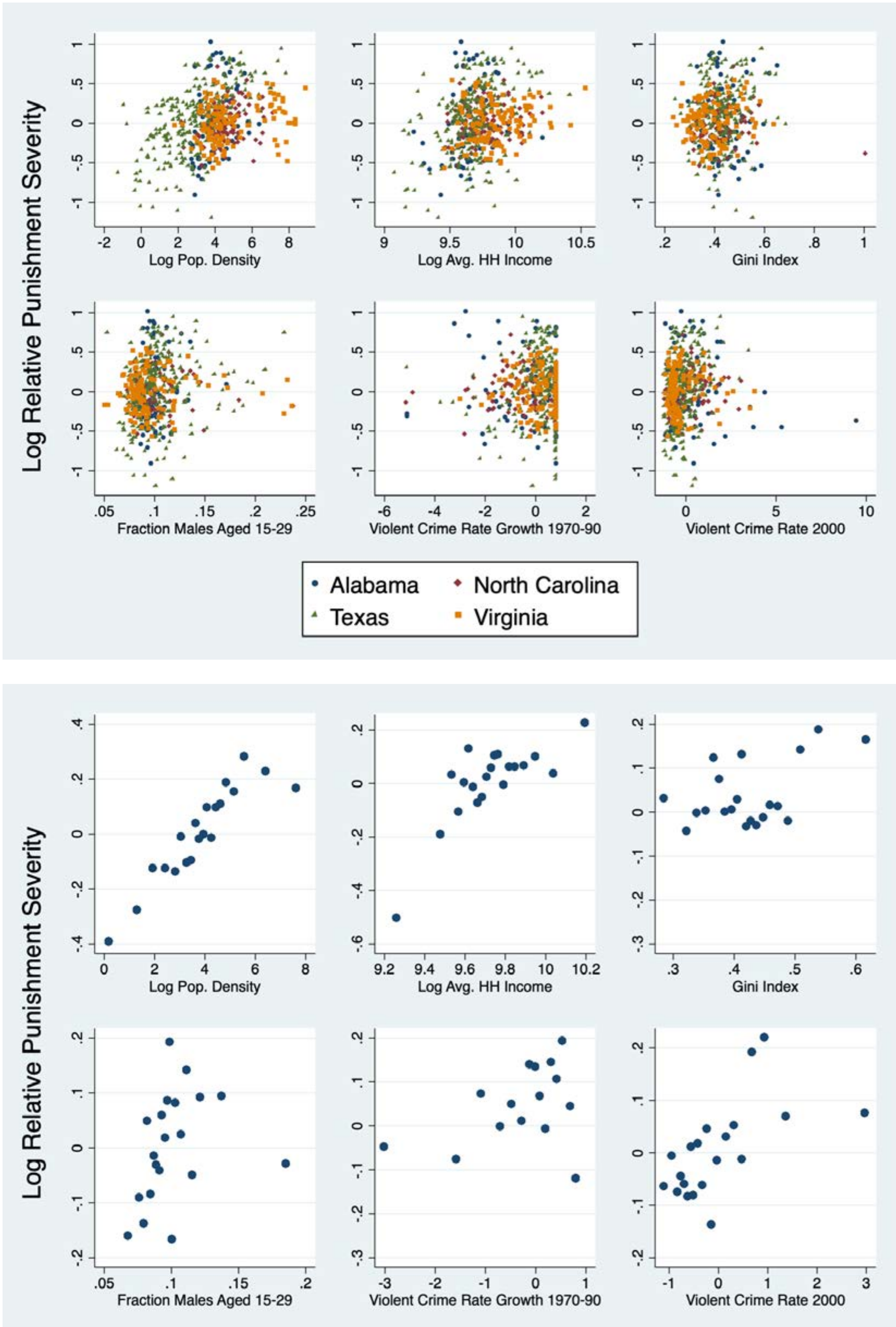
Notes: The first row reports an estimate for the variance of $\hat{y}_j$, constructed as described in Section C.3, where the outcome is confinement rate measured in percentage points. The second, third, and fourth rows report the variance of $\hat{\theta}_j$, $\hat{\gamma}_j$, and $\hat{\tau}_j$ using a split-sample approach to correct for the (correlated) measurement error in each term. The fifth, sixth, and seventh row reports the covariance between $\hat{\theta}_j$ and $\hat{\gamma}_j$, $\hat{\theta}_j$ and $\hat{\tau}_j$, and $\hat{\gamma}_j$ and $\hat{\tau}_j$, all of which is also estimated using a split-sample approach. The eighth row reports the correlation between $\hat{\theta}_j$ and $\hat{\gamma}_j$, which is also estimated using a split-sample approach.

Figure A1: Mover Defendant Event Studies, Placebo Sample



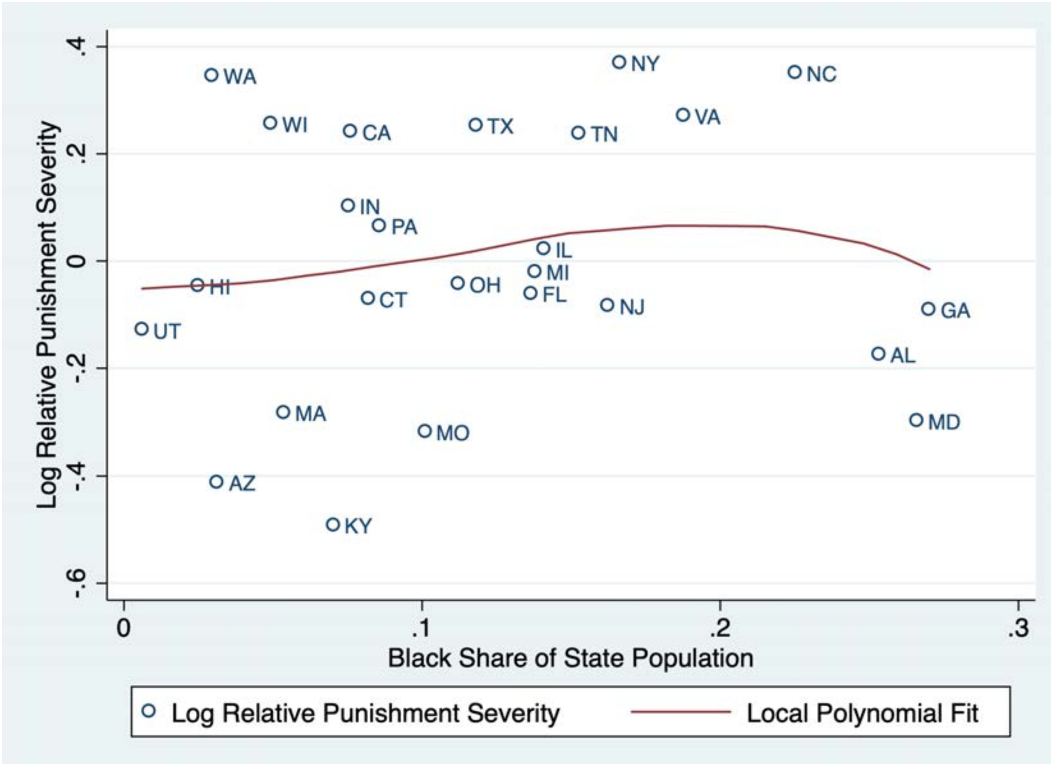
Note: We plot adjusted realized changes in confinement rate before and after the move against predicted changes by state, adjusting for offense and criminal history, and pooling by origin and destination punishment severity quartile. Marker size is proportional to the number of charges represented in the origin quartile by destination quartile by state cell. The dashed line is the $45°$ line, while the solid line is a fitted line through the points, weighted by cell size. Movers are restricted to placebo sample, defendants that are charged in multiple cases in one jurisdiction prior to move.

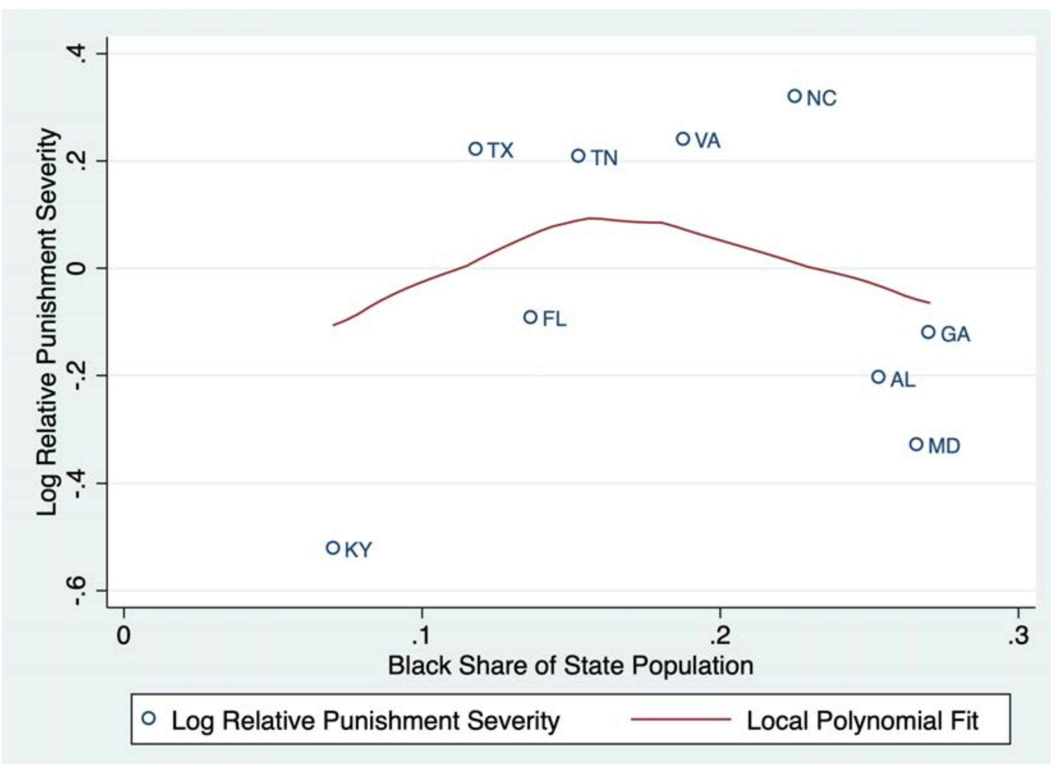
Figure A2: Punishment Severity and Jurisdiction Covariates



Note: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. In Panel A, each mark represents a jurisdiction. Panel B is a bin scatter plot where we group jurisdictions into ventiles by their covariate value.

Figure A3: Punishment Severity and Racial Heterogeneity: State Court Processing Statistics Data

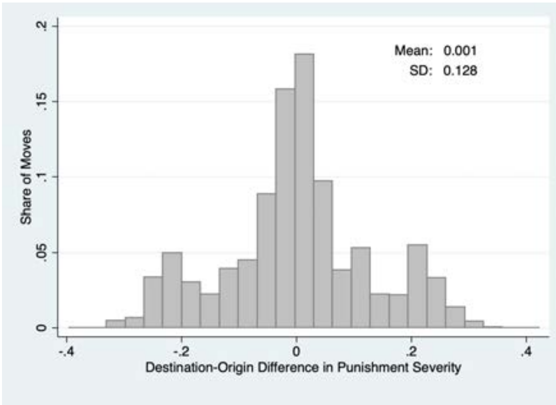




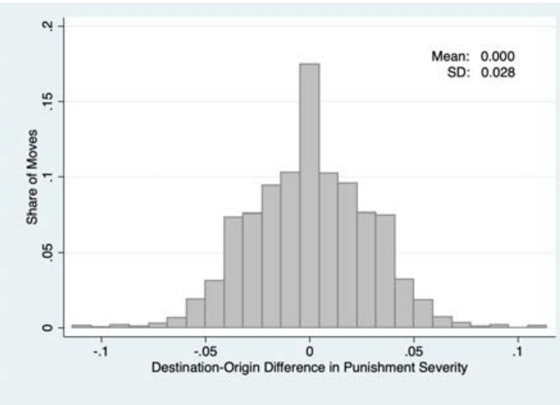
Note: Log relative punishment severity is constructed by first calculating the predicted confinement rate for each jurisdiction based on the overall composition of cases within the full sample and taking the state level average. We then divide by the overall (full sample) confinement rate and take the log of this ratio. The top panel shows all states included in the State Court Processing Statistics Data series (demeaned by region) and the bottom panel includes only Southern states.

Figure C1: Distribution of Destination-Origin Difference in Punishment Severity

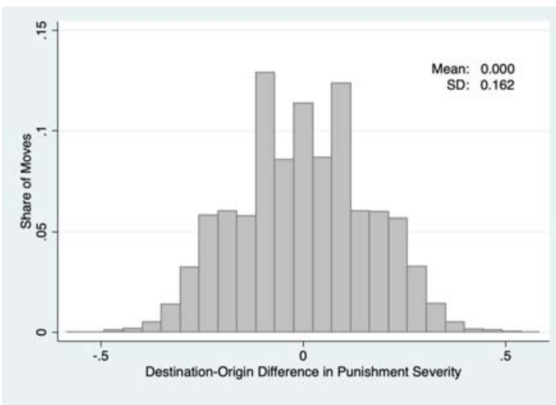
(a) Alabama



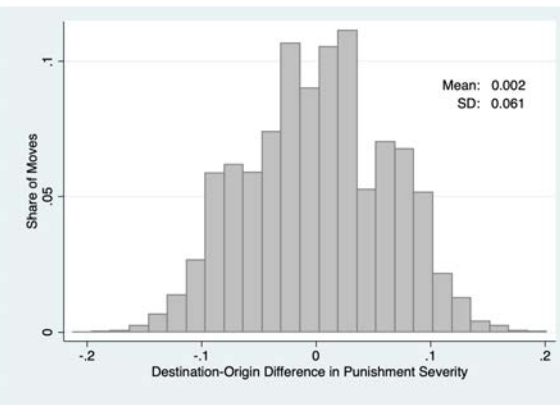
(b) North Carolina



(c) Texas



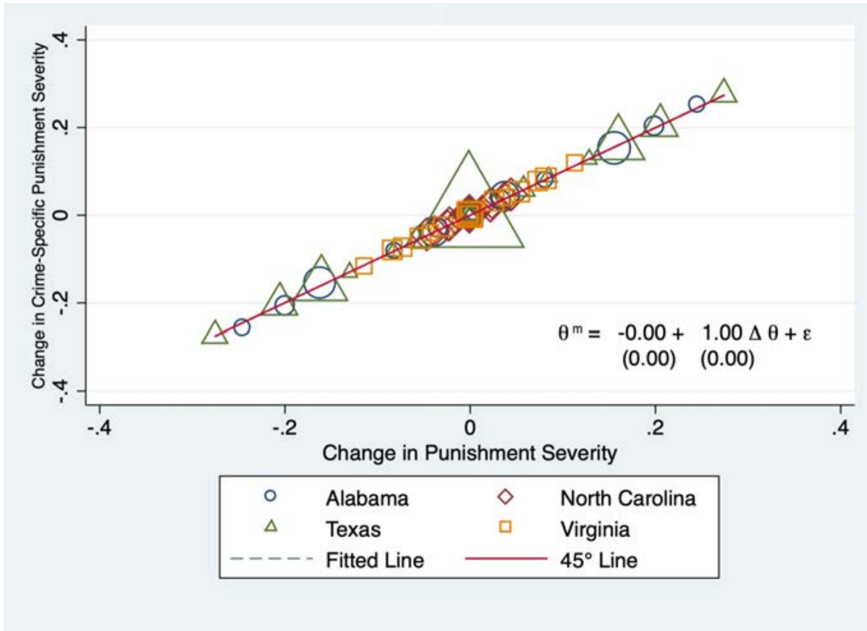
(d) Virginia



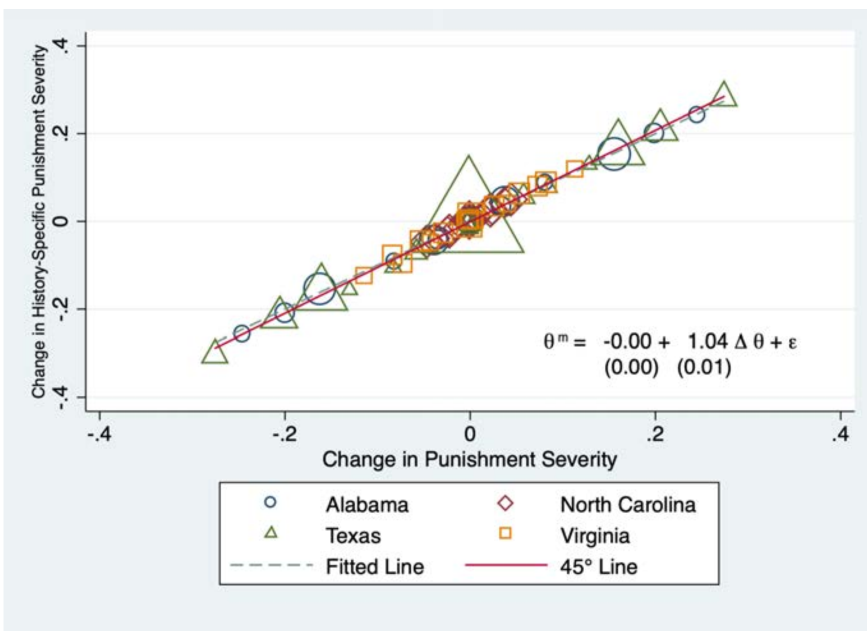
Note: We plot histograms for the difference between punishment severity θ_j in origin and destination jurisdictions among 'mover' defendants, those charged in multiple jurisdictions. We limit movers to defendants whose pre-'move' offense occurs at least two years prior to the end of the data to avoid selecting on initial sentence length. Mover defendants are described in more detail in Section 3.2.

Figure C2: Do Defendants Sort on Match Effects?

(a) By Crime Type



(b) By Criminal History



Note: In Panel A, we plot changes in jurisdiction by crime type (property, violent, drug, other) match effects against changes in punishment severity before and after the move, pooling by origin and destination punishment severity quartile. In Panel B, we plot changes in jurisdiction by criminal history category match effects against changes in punishment severity before and after the move, pooling by origin and destination punishment severity quartile. For each state, using state-specific criminal history scores, we calculate the median criminal history among those with any criminal history. We then divide defendants into three groups: those with zero criminal history, those with criminal history below the conditional median, and those with criminal history above the conditional median. In both panels, marker size is proportional to the number of charges represented in the origin quartile by destination quartile by state cell. The dashed lines are $45°$ lines, while the solid line is a fitted line through the points, weighted by cell size.