ONLINE APPENDIX

RACIAL DIVISIONS AND CRIMINAL JUSTICE: EVIDENCE FROM SOUTHERN STATE COURTS

Benjamin Feigenberg
Conrad Miller

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 iden-

tify 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.[1] 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.[2] 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

[1] This rotation has occasionally been suspended due to budget constraints.

[2] The method of election was changed to partisan elections in 2017.

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 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).

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.

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. As noted, we drop offenses that by statute cannot lead to an incarceration sentence. Based on Sec. 12.23 of the Texas Penal Code, we thus exclude Class C misdemeanors, which represent the lowest level criminal offense and do not have jail or prison penalties.

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.

As noted, we also drop offenses that by statute cannot lead to an incarceration sentence. Based on 18.2-11 in the Code of Virginia, we thus exclude Class 3 and 4 misdemeanors, which represent the lowest level criminal offenses and do not have jail or prison penalties.

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.

B.5 Federal System

As a robustness check, we alternatively define criminal history based on federal statute and verify that results are robust to doing so. 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.

C Appendix: Punishment Severity Robustness Checks

C.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 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 Appendix Table C1. 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, Heining and Kline (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 show that the patterns in punishment severity that we document are quantitatively similar for each subcategory of charges.

C.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 selection 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 (of Justice. Federal Bureau of Investigation, 2006*a*,*b*,*c*,*c*,*d*,*e*, 2007, 2008, 2009, 2011, 2014*a*,*b*,*c*,*d*, 2017*a*,*b*). 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.

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 Appendix Table C1, 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.

D Appendix: Within-Defendant Validation and Decomposition Exercises

For the exercise described below, we limit our analysis of MJ defendants to those whose pre-move offense occurs at least two years prior to the end of the data to avoid selecting on initial sentence length.

D.1 Constructing Predictions in Multi-Jurisdiction Defendant Validation Exercise

Formally, we take first-differences of equation (1) to model the *change* in charge outcomes for a defendant arrested in jurisdiction A for charge c' at time t' and subsequently arrested in jurisdiction B for charge c at time t[3]:

$$\begin{aligned} Y_{ict} - Y_{ic't'} =& (\tau_{cth(i,t)} - \tau_{c't'h(i,t')}) + (X_i\gamma^X - X_i\gamma^X) + (Z_{it}\gamma^Z - Z_{it'}\gamma^Z) \\ &+ (\theta_{j(i,c,t)} - \theta_{j(i,c',t')}) + (\epsilon_{ict} - \epsilon_{ic't'}). \end{aligned} \tag{D.1}$$

We abuse notation and write $f(i,c,t) - f(i,c',t')$ as $\Delta_i f(i,c,t)$ so that equation (D.1) reduces to

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

Appendix Figure C1 plots the distribution of $\Delta_i\theta_{j(i,c,t)}$ for multi-jurisdiction defendants, separately by state. For each defendant i we plug in values 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 multi-jurisdiction defendants:

[3] We restrict to pairs t' and t such that defendant i is not arrested between those times. That is, we limit to pairs of arrests that immediately follow one another.

$$\underbrace{\Delta Y_{ict} - \Delta\widehat{\tau_{cth(i,t)}} - \Delta Z_{it}\widehat{\gamma^Z}}_{\text{adjusted change in confinement}} = \alpha + \beta \underbrace{\Delta\widehat{\theta_{j(i,t)}}}_{\text{change in punishment severity}} + \xi_{ict} \tag{D.2}$$

adding a constant term α to allow for systematic prediction error. A β coefficient of one indicates that the punishment severity estimates provide unbiased forecasts for within-defendant changes in outcomes.

In Panel A of Figure 2, we plot actual changes in outcomes against forecasted changes separately by state, pooling by origin and destination punishment severity quartile. This can be reframed as visual IV corresponding to the two-stage least squares (2SLS) system where (D.2) is the second stage and the first stage is:

$$\Delta\widehat{\theta_{j(i,t)}} = \kappa + \pi_{Q(j(i,t')),Q(j(i,t))} + \nu_{ict} \tag{D.3}$$

where $Q(j)$ is the punishment severity quartile for jurisdiction j and $\pi_{Q(j(i,t')),Q(j(i,t))}$ are indicators for the combination of origin and destination punishment severity quartiles for defendant i first arrested in jurisdiction $j(i,t')$ and then arrested in jurisdiction $j(i,t)$. As in Angrist et al. (2017), the overidentifcation test for the system (D.2) and (D.3) measures whether the punishment severity estimates have the same predictive validity for every combination of moves. This is equivalent to testing the null hypothesis that, absent sampling error, the data points in Panel A of Figure 2 would fall on the 45° line.

In practice, the data points deviate sufficiently from the 45° line that we reject the null hypothesis in the overidentification test, which implies that our punishment severity estimates do not have the same predictive validity for every group of moves (Angrist et al., 2017). Yet the deviations are small and, reassuringly, the results we document for the relationship between local punishment severity and racial heterogeneity are unchanged if we use punishment severity estimates derived from a variant of equation (1) that includes defendant fixed effects.

D.2 Do Multi-Jurisdiction 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

$$y_{ict} = \tau_{cth(i,t)} + x_i\gamma^x + z_{it}\gamma^z + \theta^M_{j(i,c,t),k(i,c,t)} + \epsilon_{ict} \tag{D.4}$$

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 multi-jurisdiction 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.

D.3 Decomposing Punishment Severity

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

$$y_{ict} = \tau_{cth(i,t)} + \gamma_i + \theta_{j(i,c,t)} + \epsilon_{ict} \tag{D.5}$$

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, Heining and Kline, 2013). Fortunately, within each state the connected set includes all jurisdictions. We summarize punishment severity estimates derived from this approach in Panel A of Table 5.

We next construct 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. We report the results of this additive decomposition in Panel B of Table 5.

D.4 Across-Jurisdiction Variance Decomposition

We next decompose cross-jurisdiction variation of confinement rates into the variances and covariances of θ_j, $\bar{\gamma}_j$, and $\bar{\tau}_j$. Motivated by the fact that

$$Var(\bar{y}_j) = Var(\theta_j) + Var(\bar{\gamma}_j) + Var(\bar{\tau}_j) + 2Cov(\theta_j, \bar{\gamma}_j) + 2Cov(\theta_j, \bar{\tau}_j) + 2Cov(\bar{\gamma}_j, \bar{\tau}_j), \quad \text{(D.6)}$$

we compute and report the sample analog for each term. In estimating each variance and covariance term, we follow Finkelstein, Gentzkow and Williams (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 (D.5) 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 (D.6). 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 varia-

tion in confinement rates.

Table D1 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.

E Appendix: Additional Data Sources

E.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.

E.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 of Justice. Office of Justice Programs. Bureau of Justice Statistics (2014). 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.

References

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *Quarterly Journal of Economics*, 132(2): 871–919.

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

Cattaneo, Matias, Richard Crump, Max Farrell, and Yingjie Feng. 2019. "On Binscatter." Unpublished manuscript.

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

United States Department of Justice. Federal Bureau of Investigation. 2006*a*. "[Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2000." *https://doi.org/10.3886/ICPSR03451.v4*.

United States Department of Justice. Federal Bureau of Investigation. 2006*b*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2001." *https://doi.org/10.3886/ICPSR03721.v2*.

United States Department of Justice. Federal Bureau of Investigation. 2006*c*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2002." *https://doi.org/10.3886/ICPSR04009.v2*.

United States Department of Justice. Federal Bureau of Investigation. 2006*d*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2003." *https://doi.org/10.3886/ICPSR04360.v2*.

United States Department of Justice. Federal Bureau of Investigation. 2006*e*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2004." *https://doi.org/10.3886/ICPSR04466.v1*.

United States Department of Justice. Federal Bureau of Investigation. 2007. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2005." *https://doi.org/10.3886/ICPSR04717.v1*.

United States Department of Justice. Federal Bureau of Investigation. 2008. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2006." *https://doi.org/10.3886/ICPSR23780.v1*.

United States Department of Justice. Federal Bureau of Investigation. 2009. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2007." *https://doi.org/10.3886/ICPSR25114.v1*.

United States Department of Justice. Federal Bureau of Investigation. 2011. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2008." *https://doi.org/10.3886/ICPSR27644.v1*.

United States Department of Justice. Federal Bureau of Investigation. 2014*a*. "State Court Processing Statistics, 1990-2009: Felony Defendants in Large Urban Counties." *https://doi.org/10.3886/ICPSR02038.v5*.

United States Department of Justice. Federal Bureau of Investigation. 2014*b*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2009." *https://doi.org/10.3886/ICPSR30763.v2*.

United States Department of Justice. Federal Bureau of Investigation. 2014*c*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2010." *https://doi.org/10.3886/ICPSR33523.v2*.

United States Department of Justice. Federal Bureau of Investigation. 2014*d*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2011." *https://doi.org/10.3886/ICPSR34582.v2*.

United States Department of Justice. Federal Bureau of Investigation. 2014*e*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2012." *https://doi.org/10.3886/ICPSR35019.v1*.

United States Department of Justice. Federal Bureau of Investigation. 2017*a*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2013." *https://doi.org/10.3886/ICPSR36117.v2*.

United States Department of Justice. Federal Bureau of Investigation. 2017*b*. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2014." *https://doi.org/10.3886/ICPSR36399.v2*.

Table A1: Charges per Reported Crime Across Jurisdictions

	Alabama	North Carolina	Texas	Virginia
Charges per Crime UCR Part I:				
Mean	0.245	0.564	0.204	0.474
SD	(0.204)	(0.466)	(0.180)	(0.192)
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: Single-Jurisdiction ('SJ') versus Multi-Jurisdiction ('MJ') Defendants

State:	Alabama			North Carolina			Texas			Virginia		
	All	Multiple Cases		All	Multiple Cases		All	Multiple Cases		All	Multiple Cases	
		SJ	MJ		SJ	MJ		SJ	MJ		SJ	MJ
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	
	SJ	MJ	SJ	MJ	SJ	MJ	SJ	MJ
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 multi-jurisdiction ('MJ') and single-jurisdiction ('SJ') defendants as described in Section 3.2. Comparisons are based on pre- and post-move case pairs for MJ defendants and on sequential case pairs for SJ defendants. 'Move' refers to a change in the jurisdiction in which a defendant is arrested.

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 (Standardized)				0.015	0.016	0.000	0.017	0.001
				(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)	
Black Share at "Peak" Severity					0.36**	0.29**		0.26**
					(0.056)	(0.045)		(0.042)
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. In each column, Black Share at "Peak" Severity is estimated from the corresponding quadratic term coefficients on Black Share of Population. Corresponding standard errors are constructed using the delta method.
'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)	
Slave Share at "Peak" Severity	0.37**	0.42**	0.43**	0.44**	0.38**
	(0.028)	(0.057)	(0.064)	(0.078)	(0.050)
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. In each column, Slave Share at "Peak" Severity is estimated from the corresponding quadratic term coefficients on Slave Share. Corresponding standard errors are constructed using the delta method.

'†' 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)
Black Share at "Peak" Severity	0.41**	0.37**	0.47**	0.37**	0.36**	0.32**	0.30**	0.28**
	(0.103)	(0.043)	(0.104)	(0.051)	(0.062)	(0.037)	(0.043)	(0.048)
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 (D.5) (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. In each column, Black Share at "Peak" Severity is estimated from the corresponding quadratic term coefficients on Black Share of Population. Corresponding standard errors are constructed using the delta method.
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)	
Black Share at "Peak" Severity	0.30**	0.29**	0.31**	0.34**	0.29**	0.44**	0.39**	0.40**	0.42**	0.42**
	(0.020)	(0.036)	(0.041)	(0.050)	(0.034)	(0.023)	(0.023)	(0.028)	(0.037)	(0.040)
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. In each column, Black Share at "Peak" Severity is estimated from the corresponding quadratic term coefficients on Black Share of Population/Defendants. Corresponding standard errors are constructed using the delta method.
'†' 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. In practice, we use the command *psacalc* in Stata to implement relevant calculations.

'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**	1.898*	1.985*	1.772*					
		(0.851)	(0.810)	(0.815)	(0.823)					
Black Share of Population, Squared		-3.902*	-2.208	-2.023	-1.719					
		(1.591)	(1.448)	(1.449)	(1.453)					
Black Share of Defendants						2.641**	1.367˜	1.289˜	1.346˜	
						(0.708)	(0.697)	(0.718)	(0.752)	
Black Share of Defendants, Squared						-2.591*	-1.221	-0.899	-1.216	
						(1.029)	(1.009)	(1.075)	(1.132)	
Log Population Density			0.087	†	x		0.078	†	x	
			(0.058)				(0.058)			
Log Average Household Income			0.710*	0.665˜	x		0.627˜	0.587	x	
			(0.387)	(0.387)			(0.379)	(0.381)		
Gini Coefficient			-2.197**	-2.127**	x		-2.095**	-2.036*	x	
			(0.795)	(0.808)			(0.810)	(0.818)		
Fraction Males Aged 15-29			-1.478	-1.505	x		-1.419	-1.414	x	
			(2.154)	(2.225)			(2.124)	(2.199)		
Violent Crime Rate Growth, 1970-1990			-0.013	-0.018			-0.012	-0.015		
			(0.046)	(0.045)			(0.047)	(0.047)		
Violent Crime Rate, 2000			0.169**	0.170**	x		0.171**	0.173**	x	
			(0.055)	(0.054)			(0.054)	(0.054)		
Republican Vote Share, 2000										0.730**
										(0.224)
State FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.009	0.021	0.088	0.095	0.110	0.033	0.086	0.093	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 E. 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*	*0.005*	*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*	*0.011*	*0.013*	*0.004*	*0.026*	*0.034*
Black Share of County Population		-0.579	-0.997		1.471	2.973
P-value		*0.468*	*0.241*		*0.492*	*0.356*
Black Share of County Population, Squared		0.164	1.110		-1.980	-3.886
P-value		*0.913*	*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 E.

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: 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

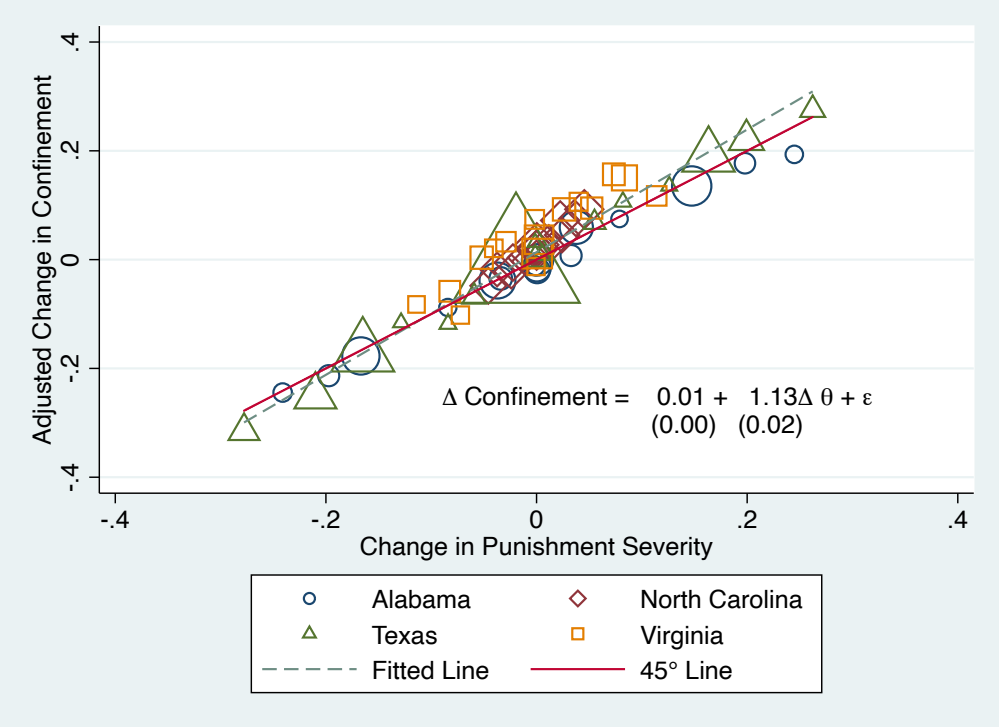
Notes: In the top panel, jurisdiction-specific punishment severity is constructed separately for each referenced subsample of defendants or charges. There are two jurisdictions where we are unable to calculate punishment severity for black defendants due to insufficient data (King County, Texas and McMullen 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.

Table D1: 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 D.4, 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.

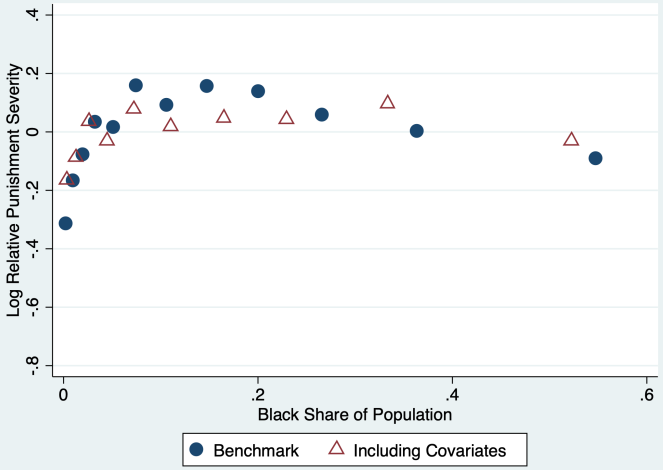
Figure A1: Forecasted and Actual Change in Outcomes for Multi-Jurisdiction Defendants, Placebo Sample



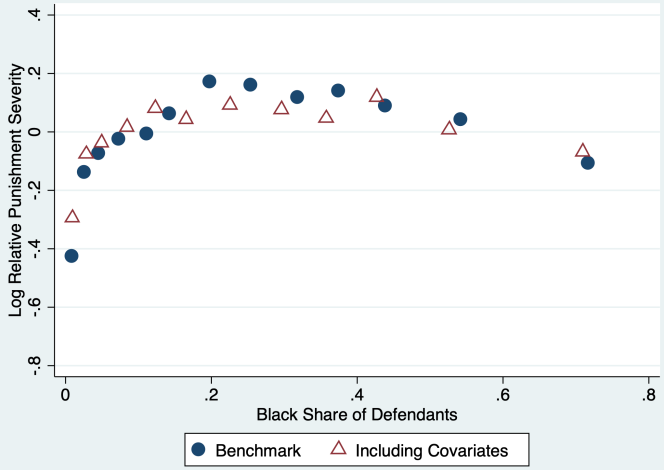
Note: We plot adjusted realized changes in confinement rate before and after the move against forecasted changes by state, adjusting for offense and criminal history, and pooling by origin and destination punishment severity quartile. Marker size is proportional to the number of charges represented in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

Figure A2: Punishment Severity and Racial Heterogeneity: Assessing Sensitivity to Controls

(a) Binned, Population

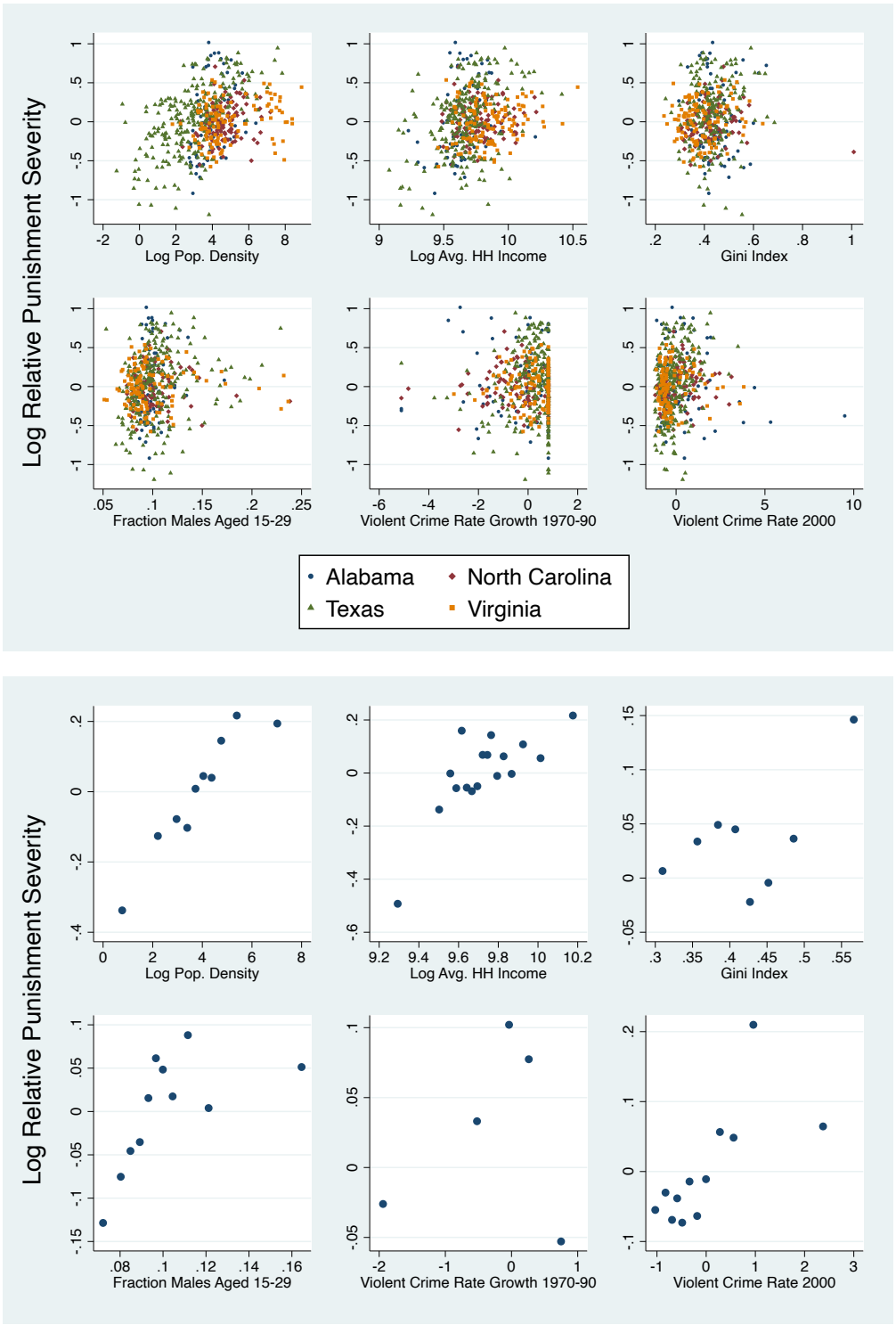


(b) Binned, Defendants



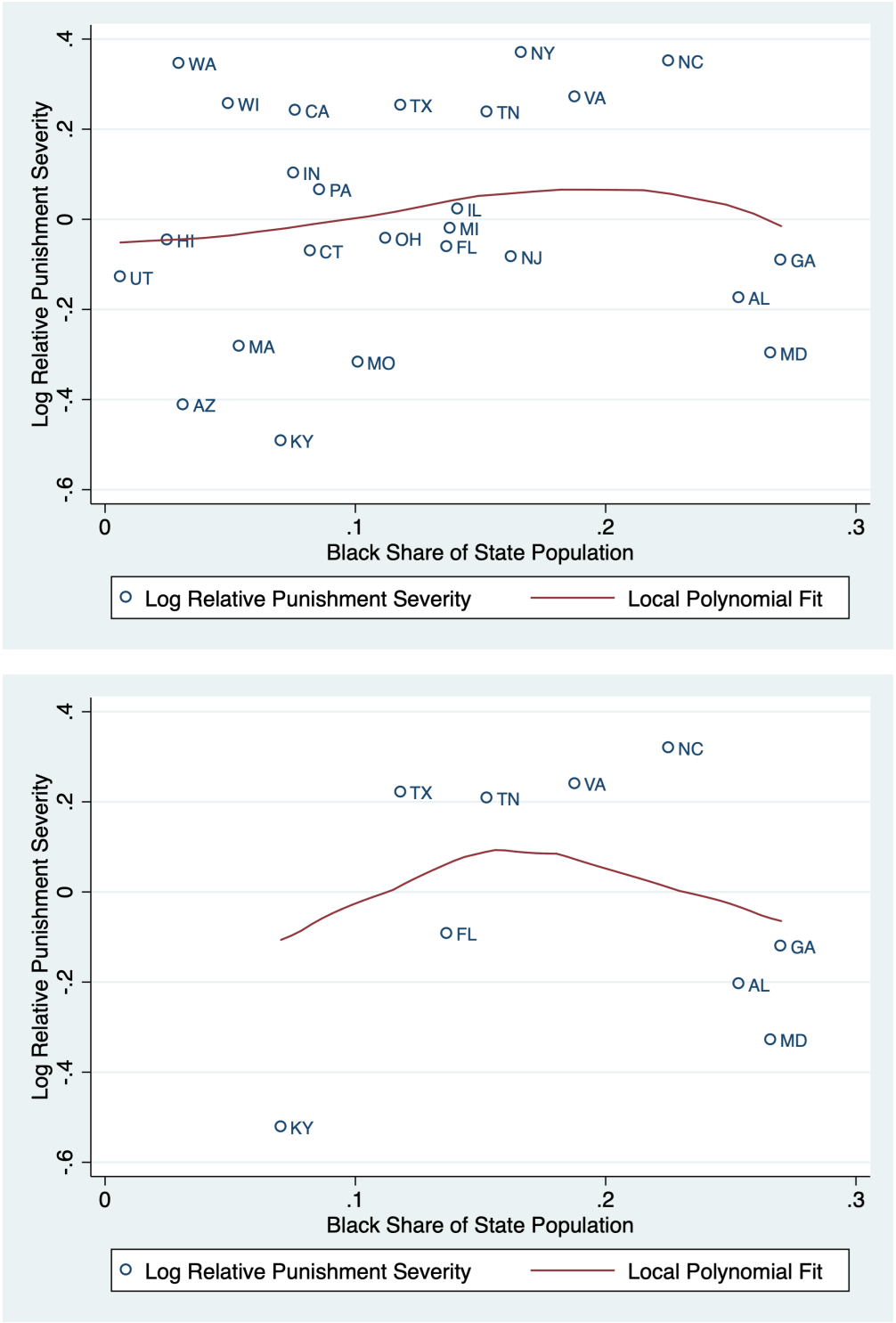
Note: Log relative punishment severity is constructed by dividing the predicted confinement rate for each jurisdiction based on the overall composition of charges within the state by the overall state confinement rate and then taking the log of this ratio. In Panels A-B, we overlay the benchmark plots presented in Panels C-D of Figure 3 with plots that adjust for the following covariates using the approach developed in Cattaneo et al. (2019): Log Population Density, Log Average Household Income, Gini Coefficient, Fraction Males Aged 15-29, and Violent Crime Rate Growth (1970-1990). We demean each covariate so that the range of y-axis values is comparable with and without covariates included.

Figure A3: 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 using the data-driven approach developed in Cattaneo et al. (2019).

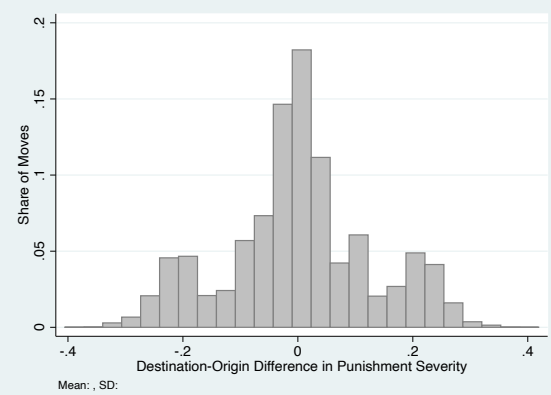
Figure A4: Punishment Severity and Racial Heterogeneity: State Court Processing Statistics Data



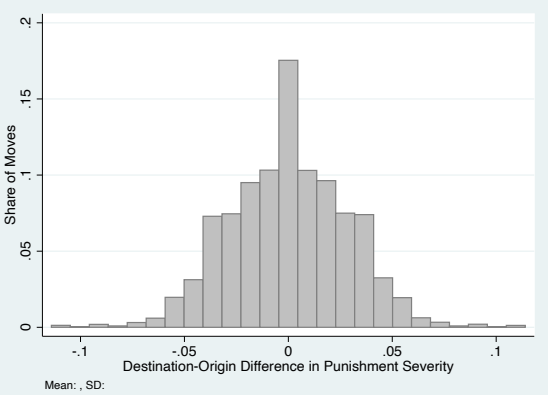
Note: Log relative punishment severity is constructed by first calculating the predicted confinement rate for each jurisdiction based on the overall composition of cases within the full sample and taking the state level average. We then divide by the overall (full sample) confinement rate and take the 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

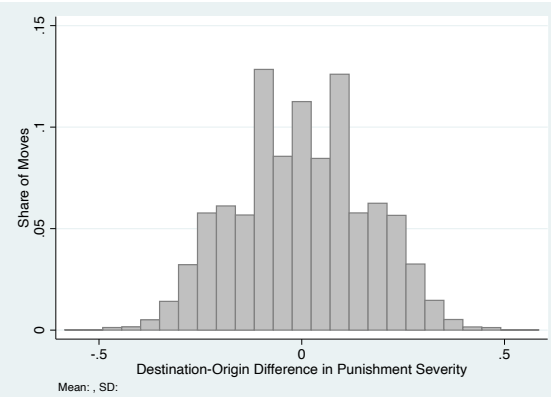
(a) Alabama



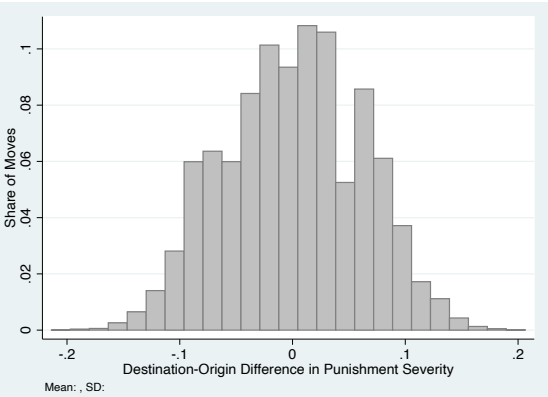
(b) North Carolina



(c) Texas



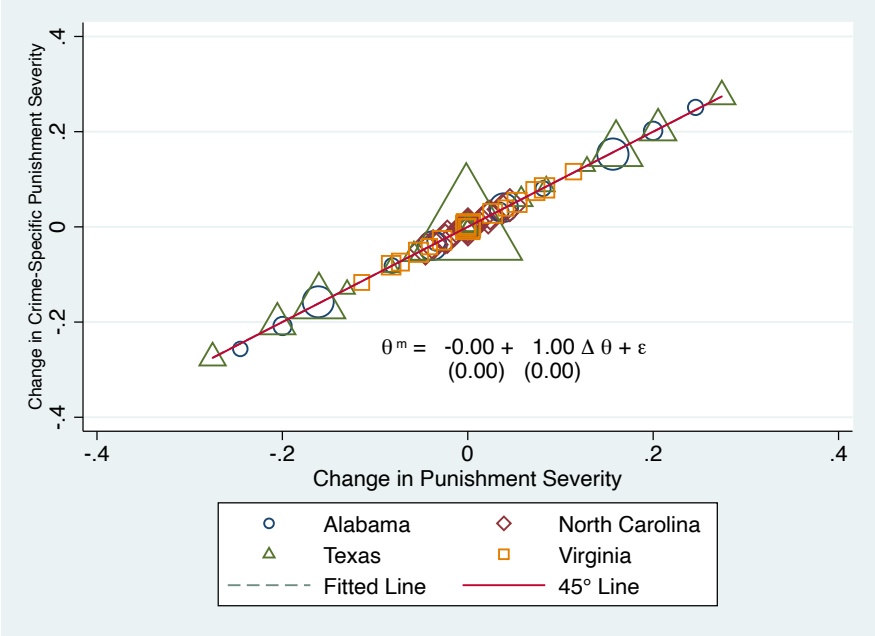
(d) Virginia



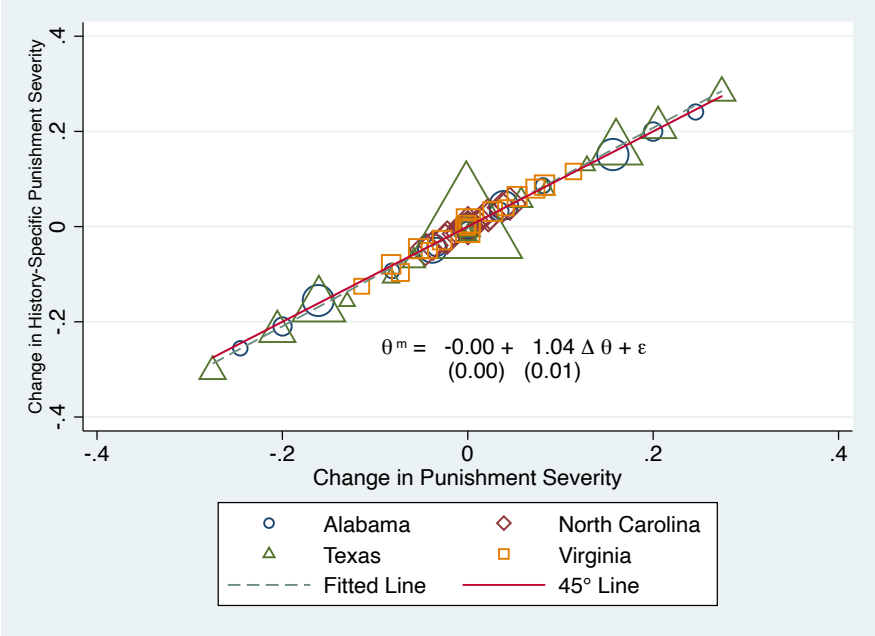
Note: We plot histograms for the difference between punishment severity θ_j in origin and destination jurisdictions among MJ defendants, those charged in multiple jurisdictions. We limit MJ defendants to those whose pre-'move' offense occurs at least two years prior to the end of the data to avoid selecting on initial sentence length. MJ 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 solid lines are 45° lines, while the dashed lines are linear best fits, weighted by cell size.