

# WHO GETS A SECOND CHANCE? EFFECTIVENESS AND EQUITY IN SUPERVISION OF CRIMINAL OFFENDERS

EVAN K. ROSE\*

December 1, 2020

Most convicted offenders serve their sentences under “community supervision” at home instead of in prison. Under supervision, however, a technical rule violation such as not paying fees can result in incarceration. Rule violations account for 25% of prison admissions nationally and are significantly more common among black offenders. I test whether technical rules are effective tools for identifying likely reoffenders and deterring crime and examine their disparate racial impacts using administrative data from North Carolina. Analysis of a 2011 reform reducing prison punishments for technical violations on probation reveals that 40% of rule breakers would go on to commit crimes if spared harsh punishment. The same reform also closed a 33% black-white gap in incarceration rates without substantially increasing the black-white reoffending gap. These effects combined imply that technical rules target riskier probationers overall, but disproportionately affect low-risk black offenders. To justify black probationers’ higher violation rate on efficiency grounds, their crimes must be roughly twice as socially costly as that of white probationers. Exploiting the repeat-spell nature of the North Carolina data, I estimate a semi-parametric competing risks model that allows me to distinguish the effects of particular types of technical rules from unobserved probationer heterogeneity. Rules related to the payment of fees and fines, which are common in many states, are ineffective in tagging likely reoffenders and drive differential impacts by race. These findings illustrate the potentially large influence of ostensibly race-neutral policies on racial disparities in the justice system.

## I INTRODUCTION

For many black men, encounters with police, courts, and prisons are as common as employment. Black high school dropouts, for example, are almost as likely to be incarcerated

\*Post-Doctoral Researcher, Microsoft Research; ekrose@gmail.com. I thank Pat Kline, David Card, Danny Yagan, and Chris Walters for their help and encouragement. I also thank the editor and four anonymous referees for their constructive feedback. This paper has benefited tremendously from comments and suggestions from Alessandra Fenizia, Fred Finan, Ingrid Haegele, Jonathan Holmes, Hilary Hoynes, Peter Jones, Nick Li, Juliana Londoño-Vélez, Maxim Massenkoff, Justin McCrary, Conrad Miller, Steven Raphael, Emmanuel Saez, Jonathan Schellenberg, Yotam Shem-Tov, Francis Wong, and seminar participants at the University of California at Berkeley. I am grateful to the North Carolina Department of Public Safety and Administrative Office of the Courts for their help in securing and understanding the data, as well as to Ginny Hevener, Linda Mitterling, George Pettigrew, Alan Pistick, Cara Stevens, and the officers of the 14th Probation District for their feedback and patience.

as to be holding a job. Recent research has studied racial disparities in decisions by police, judges, prosecutors, and juries (Fryer, 2019; Arnold et al., 2018; Rehavi and Starr, 2014; Anwar et al., 2012) and how arrest, conviction, and incarceration affect economic outcomes (Western and Pettit, 2005; Lyons and Pettit, 2011; Agan and Starr, 2018; Dobbie et al., 2018; Harding et al., 2018; Chetty et al., 2018; Bayer and Charles, 2018; Mueller-Smith and Schnepel, 2019; Bhuller et al., 2019). However, less attention has been paid to the impact of community supervision, the most common punishment in the United States. Every year, more than 4.4 million convicted offenders are sent home under supervision and on the condition that they obey strict technical rules. Breaking these rules, which often forbid alcohol and drugs, entail frequent meetings with a caseworker, and require timely payment of fees and fines levied by the court, can result in incarceration. Supervised offenders are as likely to be incarcerated for such “technical violations” as for new criminal offenses nationally (CSG, 2019), with violations particularly concentrated among black men. This “second chance” sentence is therefore a key driver of incarceration overall and of racial disparities in prison exposure.<sup>1</sup>

Technical rules, however, are the primary tools the corrections system uses to surveil supervised offenders and support their reintegration (Piehl and LoBuglio, 2005). Despite the costs, punishing technical rule breaking with incarceration—or “revoking” supervision—may therefore be effective if violations are strong indicators of future criminal behavior, making them good tags for reoffending risk, or if the threat of harsh punishments encourages compliance, which may directly benefit offenders and their communities. The *effectiveness* of revocation thus depends on how well rule violations target potential reoffenders and on any behavioral responses to potential punishments. The *equity* implications depend on racial differences in the association between rule breaking and reoffending risk (Kleinberg et al., 2017a,b) and on differences in any behavioral responses to the threat of punishment.

This paper examines the effectiveness and equity of revocation in the probation system, which accounts for 80% of the supervised population. I test whether revocation targets probationers who would otherwise commit crimes, measure its deterrence effects, and examine racial differences in targeting and deterrence. To do so, I analyze a major 2011 reform in North Carolina that reduced incarceration punishments for nonpayment of cash fees and fines, drug and alcohol use, and other rule violations. As a result, many probationers who would have been imprisoned prior to the reform instead remained in their communities or were subject to short periods of confinement. Measuring the resultant increases in arrests thus allows me to assess how effectively revocation targeted would-be reoffenders and measure any behavioral response to the change in punishments. Analyzing the reform by race allows me to examine equity in targeting and deterrence effects.

I begin with a reduced-form analysis of the 2011 reform. The analysis examines technical revocation and criminal arrests over the first year of probation for successive cohorts who started their probation spells within four years of the reform. To control for any time trends in crime, probationers are compared to individuals convicted of similar offenses and placed on unsupervised probation, an alternative punishment where technical rules are not enforced and the reform had no impact. This control group’s outcomes track the treated

1. These concerns became headline news in 2017 when the musician Meek Mill was incarcerated for breaking the terms of a decade-old sentence over technical violations that included riding a dirt bike without a helmet and traveling for performances. Jay-Z, writing in the New York Times, argued “What’s happening to Meek Mill is just one example of how our criminal justice system entraps and harasses hundreds of thousands of black people every day...Instead of a second chance, probation ends up being a land mine, with a random misstep bringing consequences greater than the crime” (Nov. 17, 2017).

group’s closely over the full pre-reform period. Results change little, however, in a simple pre-post analysis of the treated group alone.

Difference-in-differences estimates reveal that technical revocation in the first year of probation fell by 5.5 percentage points (p.p.) as a result of the reform, a 33% drop relative to the pre-reform mean of 15.4%. Arrests increased by 2.0 p.p. overall. Remarkably, the reform’s impact on black probationers’ revocation rates was nearly twice as large as its impact on non-black probationers’. As a result, racial gaps in revocation for breaking technical rules were eliminated, and thousands more black probationers were allowed to remain in their community. Yet black probationers saw only slightly larger increases in arrests after the reform than non-black probationers. The reform therefore eliminated racial gaps in revocation without impacting racial gaps in reoffending rates.

I interpret these results through a simple empirical model with two binary outcomes: whether a probationer is revoked for technical violations and whether he is arrested for a crime. Casting the reform as an instrument in the classic [Angrist et al. \(1996\)](#) framework, one can estimate the probability probationers spared revocation would reoffend instead ([Abadie, 2002](#)), which reflects the accuracy associated with using revocation as a tag for counterfactual reoffending. Combining this probability with other observed quantities one can also estimate the share of non-reoffenders who break rules (i.e., type-I error or false positives) and reoffenders who do not (i.e., type-II error or false negatives). All three concepts are relevant for the effectiveness and equity of technical rule breaking as a tag for reoffending risk.

Critically, this approach rules out any direct behavioral responses to the reform. Additional exercises support this assumption by showing that mechanical changes in revokes alone fully account for observed increases in arrests. Nevertheless, I show later that results change little when using a semi-parametric competing hazards model to relax this restriction ([Cox, 1962](#); [Tsiatis, 1975](#); [Heckman and Honoré, 1989](#); [Honoré, 1993](#); [Abbring and Van Den Berg, 2003](#)).

Applying this framework to the reduced-form results implies that roughly 37% of individuals spared revocation due to the reform were arrested instead. This estimate of the accuracy of revocation for the drug and administrative rules affected by the reform (i.e. the probability of offending conditional on revocation) is roughly 10 p.p. higher than mean arrest rates. Using revocation as a tag for reoffending risk therefore does meaningfully better than random chance. Yet both type-I and type-II errors are large, at 6% and 94%, respectively, implying revokes catch a meaningful fraction of non-reoffenders and few potential reoffenders.

The effectiveness of revocation as a tag for counterfactual risk varies substantially by race. Roughly 50% of non-black probationers spared revocation were arrested, while only 30% of black probationers did the same. The implied accuracy of revocation is therefore 66% higher in the non-black population. In fact, among black offenders accuracy is close to mean reoffending rates, implying rule breaking is no better signal of future reoffending than a coin flip. While type-II error rates are similar in both groups, type-I error rates are three times higher in the black population. Substantially more black offenders who would not have offended in the first year of their spell were therefore revoked due to technical rule violations. A decomposition exercise shows that while black offenders are more likely to reoffend overall, racial differences in reoffending risk explain less than 10% of race gaps in the likelihood of revocation, with differences in error rates explaining the remainder.

These findings are remarkably consistent with previous work in [Sakoda \(2019\)](#). Using a similar difference-in-differences strategy, [Sakoda \(2019\)](#) finds that eliminating post-release

supervision (a variation of parole) for low-risk offenders incarcerated in Kansas reduced overall rates of and racial disparities in reincarceration, but had no effects on new convictions for felony offenses. Hence racially disparate effects estimated here may reflect the impacts of technical rules used in criminal supervision systems generally rather than factors specific to North Carolina or probation.

Additional results suggest that race gaps in my setting arise due to the disparate *impact* of race-neutral rules rather than disparate *treatment* by those who enforce them. For example, there is no racial gap in punishments conditional on breaking the same rule. Moreover, technical violations for which officers have wide enforcement discretion and those that are detected automatically both exhibit large race gaps. There is also no evidence of officer-probationer race match effects. This setting thus underscores the potential importance of how rules and policies are designed rather than how they are applied for explaining racial disparities (Bushway and Forst, 2013; Neal and Rick, 2016). However, it is possible that biases undetected in these exercises remain a factor. Substantial evidence of disparate treatment for taste-based or statistical reasons (Becker, 1957; Phelps, 1972; Arrow, 1973) has been documented throughout the justice system (Abrams et al., 2012; Rehavi and Starr, 2014; Fryer, 2019; Arnold et al., 2018).

I use these results to conduct a partial cost-benefit analysis that compares the costs of revoking a rule breaker to the social costs of any arrests and attendant punishments if allowed to remain free. The results show that for every \$100 the state spends revoking rule breakers, it saves \$30 it would have spent on incarceration for new arrests. To justify the state’s use of revocation, the social costs of arrests averted must fill the gap, implying a break-even valuation of roughly \$40,000 per arrest. Because black probationers are targeted more aggressively, break-even valuations for black offenders are roughly twice as large as for non-black offenders. Using estimates from the existing literature, I find that the social cost of averted offenses may fall below this benchmark, although estimates are noisy.

While this quasi-experimental evidence is informative, several important issues are more difficult to address. First, the timing of arrests and rule violations are potentially crucial drivers of effectiveness and any disparate impacts. For example, if all arrests happen early in spells but all rule violations happen later, rules are unlikely to be useful for incapacitating reoffenders even when the propensities to reoffend and break rules are tightly correlated. Second, probationers may change their behavior in response to changes in rule enforcement, an effect ruled out in the empirical model. And finally, different types of rules may have very different impacts. The reform however, affected multiple rules simultaneously, making it difficult to use the quasi-experimental variation to estimate the accuracy and error rates of, for example, drug or fees and fines violations specifically.

I address these questions using a semi-parametric model of competing hazards. The model endows probationers with latent risks of rearrest and revocation that allow for state dependence and depend on observables such as age and criminal history and on unobserved random effects. The multiple-spell nature of my data allows me to flexibly model the distribution of this unobserved heterogeneity and its correlation across risks (Heckman and Honoré, 1989; Honoré, 1993; Abbring and Van Den Berg, 2003). Each risk can also shift in response to the 2011 reform to directly capture any behavioral responses to the change in policy. An extension disaggregates rule types by breaking the risk of revocation into type-specific risks that also depend on observed and unobserved factors. By estimating the model completely separately by race and gender, I can therefore capture rich differences in the relationship between rule breaking, reoffending, and enforcement across populations.

The estimates show that revocation and arrest risk are tightly correlated, but much less

so for black offenders. Black probationers who would not be rearrested within three years are roughly 60% more likely to be revoked for rule breaking than comparable non-black offenders. Estimated behavioral responses to the change in policy are also small. Weekly average latent arrest hazards are less than 0.1 p.p. higher after the reform and the risk of drug violations and failure to pay fees and fines decreases slightly. Though perhaps surprising, these limited behavioral responses are consistent with a series of randomized controlled trials showing that intensive monitoring and more stringent supervision conditions typically fail to impact probationers’ behaviors (Hennigan et al., 2010; Barnes et al., 2012; Boyle et al., 2013; Hyatt and Barnes, 2017).

Estimates of the impact of specific types of rules show that all rules tend to target black offenders more aggressively. However, rules related to cash fees and fines are particularly problematic. Not enforcing them would increase the share of reoffenders who break technical rules and decrease the share of non-reoffenders revoked for doing so. Hence, eliminating revocation for this type of rule provides a double social benefit by improving the effectiveness of the probation regime overall and reducing existing disparities. Since North Carolina’s 2011 reform directly addressed fees and fines, it had large impacts on disparities within more limited impacts on reoffending. Revocation for other rule types, such as drug abuse and reporting rules, tend to perform better.

Taken together, my results show how ostensibly race-neutral policies—in this case common sense rules designed to promote public safety—can generate large racial disparities not justified by the policies’ ultimate goals. In some contexts, opting to give local decision makers more discretion instead of relying on uniform rules may increase policies’ effectiveness and fairness by taking advantage of agents’ superior information and encouraging effort (Aghion and Tirole, 1997; Kuziemko, 2013; Duflo et al., 2018). North Carolina’s reform shows that holding discretion fixed, however, there is the potential to redesign rules themselves to improve outcomes. Suboptimally designed rules and policies are a potentially powerful explanation for many observed racial disparities in criminal justice, where the use of detailed guidelines to constrain decisions has become increasingly popular.

The remainder of this paper is structured as follows. I first describe the probation system both nationally and in North Carolina, explain the sources and content of my data, and estimate observational racial disparities in Section II. Section III lays out the empirical model. Section IV presents the main results that analyze the 2011 reform. Section V estimates a competing risk model for probation violations and arrests. Section VI concludes.

## II SETTING AND DATA

### *IIA The community supervision system*

Over the past several decades, the population under community supervision has grown in tandem with incarceration rates. There are now 4.4 million convicted offenders on supervision in the United States, a more than 300% increase over levels in 1980, and more than twice as many as are incarcerated today. 80% of these offenders are serving terms of probation, a period of community supervision ordered in lieu of incarceration. The remaining 20% are on parole, which is served after a period of incarceration. Since probation and parole spells can be quite short, this population turns over quickly—1.8 million individuals entered probation in 2018, and 1.9 million individuals exited (Kaeble and Alper, 2020). Many millions more individuals have thus likely served a probation or parole sentence at some point in the past. For much of the last 25 years, North Carolina operated a very small

parole system, opting to release most incarcerated individuals with no supervision. I thus focus exclusively on the probation system in this analysis.<sup>2</sup>

The size of the probation system reflects its popularity as a criminal sentence. In the 75 largest counties in the US, 51% of felony defendants receive probation as part of their sentences, with higher rates for non-violent property and drug offenders (Reaves, 2013). Misdemeanor defendants, who account for the bulk of cases processed in state courts, receive probation at even higher rates. While probation is common overall, it is used most often for young and first-time offenders facing their first serious criminal case. In North Carolina, for example, 78% of first-time felons are placed on probation, along with 70% of 16-25 year-old offenders.<sup>3</sup>

Probation spells typically last between one and three years (Reaves, 2013). Over this period, offenders must comply with a set of conditions imposed by the court as “reasonably necessary to ensure that the defendant will lead a law-abiding life or to assist him to do so” (NC General Statutes §15A-1343). Failure to comply risks incarceration for the duration of a sentence that was “suspended” at conviction. For misdemeanor offenders, suspended sentences are typically 1-5 months. For felony offenders, they range from several months to two years. North Carolina’s probation conditions include a set of standard rules: pay fees and fines ordered by the court, including general court fees of roughly \$150, a \$30-50 monthly fee for supervision itself, repayment for any indigent defense provided (at least \$60), and any restitution ordered (Markham, 2018),<sup>4</sup> remain within the jurisdiction of the court unless given permission to travel, report regularly to a probation officer, submit to drug and alcohol tests and warrantless searches, and attempt to remain gainfully employed. Occasionally, judges impose special conditions such as substance abuse treatment programs and electronic monitoring.<sup>5</sup>

These technical rules and requirements serve two main purposes (Piehl and LoBuglio, 2005). First, they are intended to help offenders successfully reintegrate and support rehabilitation. Second, they serve as an “early warning” system that allows the corrections system to preempt potentially serious criminal offending. As is clear from North Carolina’s statute, this public safety motive is an important rationale for enforcing technical rules. Interviews conducted with probation officials, probationers, judges, and attorneys across the country by the University of Minnesota’s Robina Institute show that many other jurisdictions have a similar focus (Robina Institute, 2016).

North Carolina operates two forms of probation: supervised and unsupervised. Supervised probationers are assigned a probation officer who is personally responsible for monitoring them. These officers oversee 60-80 offenders at a time, conducting regular interviews, drug tests, searches, and arrests. Most officers have four-year degrees in a criminal justice related field. Roughly 50% of officers are female and 40% are black. Unsupervised probationers are not assigned a probation officer. They are technically subject to the same rules as their supervised peers, except those related to supervision, such as reporting regularly to an officer. While in some cases judges have discretion to assign either supervised or un-

2. Sakoda (2019) studies the impact of a variant of parole called post-release supervision in Kansas.

3. Individuals granted deferred prosecution are also typically placed on probation. Unlike regular probationers, however, after successfully completing their spell their records may be cleared.

4. There is no data available on collection rates for court costs in North Carolina. Surveys in other districts have found overall repayment rates ranging from 50% to 9% in other states (Pepin, 2016).

5. The full set of regular and special probation conditions are listed in North Carolina’s general statutes, available at: [https://www.ncleg.net/EnactedLegislation/Statutes/PDF/ByArticle/Chapter\\_15A/Article\\_82.pdf](https://www.ncleg.net/EnactedLegislation/Statutes/PDF/ByArticle/Chapter_15A/Article_82.pdf). Unfortunately, high quality data on total fees and fines assessed is difficult to obtain. Existing research suggests total fines can easily breach \$1,000 (Hunt and Nichol, 2017).

supervised probation, unsupervised probation tends to be reserved for misdemeanants and individuals convicted of driving while intoxicated or with a revoked license. Due to the lack of monitoring, unsupervised probationers are rarely subject to technical rule violations and thus were largely unaffected by North Carolina’s 2011 reform, making them a useful control group.

Probation officers’ responses to “non-compliance” are guided by a detailed grid that specifies appropriate responses as a function of the offender’s risk level and behavior. Non-willful violations are typically dealt with via formal reprimands or further investigation. When an officer detects a willful violation of a probation rule, she initiates the violation process by filing a formal report.<sup>6</sup> The offender must then report to a local judge for a violation hearing. Judges can respond by “revoking” probation and sending the individual to jail or prison for the duration of their original, suspended sentence. Judges can also modify specific conditions, extend the supervision term, and issue verbal reprimands and warnings. In practice, judges closely follow probation officers’ recommendations, agreeing to revoke in 85% of hearings where the officer favors doing so, and revoking in 45% of hearings overall. Over the 2000s, probationers incarcerated without a new criminal conviction accounted for ~40% of NC state prison admissions.

Probation officers may have some discretion in whether to code a behavior such as drug use as a technical or criminal violation. Throughout the analysis, I define a technical revocation as a revocation without an intervening arrest by regular North Carolina law enforcement. Although most probation violations for new criminal behavior are accompanied by a new criminal arrest, occasionally they are not. This definition thus avoids relying on violation codes themselves to define rule-driven incarceration, which is attractive because violation coding may vary across groups or be affected by the reform. Estimated racial disparities change little, however, when using alternative definitions of technical revocation, such as revocation for violations coded by officers as non-criminal.

## *IIB 2011 reform*

In 2011, North Carolina made major changes to the state’s criminal justice system by passing the Justice Reinvestment Act (JRA).<sup>7</sup> Among the most consequential changes was the introduction of strong limits on courts’ authority to revoke probation. For all probation violations occurring on or after December 1, 2011, supervision could be revoked only for new criminal offenses or for absconding (i.e., fleeing supervision).<sup>8</sup> Previously, judges could revoke for any technical violation, including non payment of fees and fines, not reporting, or failing drug and alcohol tests.

6. I shadowed probation officers at work in Durham, N.C. for several days during the summer of 2018. Officers rely heavily on their forms and computer systems. They are primarily incentivized to ensure that all appropriate policies and procedures are followed in each case. Many interactions with offenders consist of probation officers clicking through automated forms on their desktop computers while the probationer answers a standard set of questions. Most officers described their responsibilities as ensuring that their caseload respects all conditions imposed in their sentences, not helping to identify and incapacitate the riskiest offenders.

7. The law reflected several years of work by the Council of State Governments’ Justice Center (CSG). After studying North Carolina’s corrections system, the CSG concluded that rule-driven incarceration was responsible for hundreds of millions of dollars of corrections costs annually (*alias?*). Law makers passed the JRA in an effort to reduce these costs and lower projected correctional spending in the future.

8. The sole exception: revocation was still permissible if an offender had been punished with a short confinement spell—a CRV, discussed below—for two previous violations. Probation spells are typically not long enough for this to occur.



JRA also introduced a new violation response, called Confinement in Response to Violation (CRV), as a substitute for revocation. CRVs confine misdemeanor offenders for no more than 90 days (with duration set by the judge) in local jails and felony offenders for exactly 90 days in state prisons, with some jail credits applied in both cases. In practice, this means that some offenders no longer revoked due to the reform received a CRV instead. For felony offenders, who serve their CRVs in state prisons where I can most reliably measure incarceration, JRA’s changes amounted to a 200 day reduction on average in incarceration for technical violations in the first year of their spells.<sup>9</sup> Misdemeanor offenders typically serve their CRVs in local jails, where data on duration is not systematically collected.<sup>10</sup> In 2015, however, the North Carolina legislature eliminated CRVs completely for misdemeanor offenders (NC S.L. 2015-191), allowing me to examine how outcomes change after this alternative to revocation was removed. I return to how CRVs impact the interpretation of my results in Sections [IIC](#) and [IVD](#).

JRA also made several other changes to community supervision. Probation officers received expanded authority to impose conditions such as additional community service and “quick dips” (2-3 day jail confinements) in response to failures to comply with certain conditions. A useful feature of the reform is that changes to revocations applied to all *violations* after December 1, 2011, while changes in officers’ authority applied to probationers whose original *offenses* were committed after December 1. This allows me to study the change to revocation policy while holding officer authority constant by looking in a relatively narrow window around December 1.

Finally, JRA also made several changes to other parts of the court system, including increasing the scope of post-release supervision (a variation on parole), adjusting some sentencing enhancements, and re-defining some conditions of supervision. Since the focus of this paper is on the probation system, these changes are beyond the scope of this study.

## ***IIC Data sources***

This project primarily analyzes administrative data sets provided by the North Carolina Department of Public Safety (DPS). The core data consist of records for the universe of individuals serving supervised probation sentences that started between 2006 and 2018 (inclusive). These data detail individual demographics, the duration of spells, the original convictions, and probation officers assigned. The data also record all violations in dozens of unique categories, the probation officer’s recommended response, and the ultimate disposition. For some exercises, I categorize violations into four groups: drug related, administrative, absconding, and new crime, where absconding is the only technical violation that could be punished with revocation post-JRA.<sup>11</sup>

In addition to these records, I use data on all criminal court cases disposed from 2006 to the present provided by the Administrative Office of the Courts (AOC). Because police officers are the charging agency in North Carolina, these records capture close to the universe of arrests.<sup>12</sup> I use the AOC data to measure new criminal offenses, the length of

9. See Appendix Table [A11](#), discussed further below. The average incarceration duration for felony probation revocations in 2010 was 221 days (median 188).

10. The length of CRVs is also not recorded in the court datasets described below. Before June 2013, CRVs did not have to be served over a period of consecutive days (i.e., uninterrupted) (NC S.L. 2013-101).

11. The top violations in each category are reported in Appendix Table [A1](#).

12. In Charlotte-Mecklenberg, where I have collected jail booking records directly from the Sheriff, 93.3% of arrests appear in the AOC data. The remaining 6.7% of Charlotte records reflect non-arrest events, such as federal prison transfers.



any incarceration sentences meted out as a result, and criminal histories, and to identify individuals placed on unsupervised probation. I combine this data with additional records from the DPS that detail all sentences to supervised probation and incarceration in state prisons from the 1970s to the present.

Lastly, in some descriptive regressions I use scores on standardized, state-wide tests administered in math and reading at the end of grades three through eight. These data are housed at the North Carolina Education Research Data Center and were linked to North Carolina criminal records for related work in [Rose et al. \(2019\)](#). Test scores are only available for about a third of the sample, since not all offenders were educated in the state at times covered by the data.

All data are linked using a combination of personal and administrative identifiers. This includes full name and date of birth in all cases, but also partial social security numbers, driver’s license numbers, and unique codes assigned to individuals by the State Bureau of Investigation, Federal Bureau of Investigation, and the DPS.

### ***IID Descriptive statistics***

Descriptive statistics for the treated and control samples are provided in Table I. Both groups are young, with 50% of the sample 30 or under at the start of their spell, predominately male, and over-represent minorities relative to North Carolina’s population. Supervised probation spells last about 20 months on average and are the result of a relatively even mix of felony, misdemeanor, and driving while intoxicated or driving with a revoked license offenses. The treated sample has very limited criminal histories, with the median defendant having just one prior misdemeanor conviction and no prior sentences to supervised probation or incarceration. As expected, unsupervised probationers were convicted of less severe offenses and have more limited criminal histories.

As shown in Table II, the majority of probation spells include at least one violation, with citations for non-payment of fees and fines occurring in 50%. The next most common violation is for not reporting to a probation officer—for example by missing a weekly check-in at the local probation office. This violation occurs in 29% of spells. Drug violations and treatment program failures are also common, occurring in 18% and 16% of spells, respectively. Violations for new misdemeanor arrests are the fourth most common violation; new felony arrests are the 11th. Probationers are twice as likely to be cited for moving or changing jobs without notifying their probation officer as for committing a new felony crime.

### ***III E Racial disparities***

Racial disparities are a pervasive feature of the US criminal justice system. Black men who did not complete high school, for example, are almost as likely to be incarcerated as at work and are employed half as frequently as similarly educated white men.<sup>13</sup> Probation has been shown to contribute to these patterns ([Jannetta et al., 2014](#)). In North Carolina, black offenders face more violations of all types, as shown in Figure I. This figure reports coefficients from regressions of a black indicator on an indicator for different types of violations ever occurred within a spell. The blue bars include no additional controls, while the regressions underlying the red bars feature a battery of other factors, including demo-

13. See Appendix Figure A1.

graphics, geography, criminal history, and standardized math and reading test scores.<sup>14</sup> The first blue bar, for example, shows that black probationers are 17 p.p. more likely to face administrative violations, a 30% increase relative to the non-black mean. After including all controls, this difference drops to about 10 p.p. In all cases, however, the black coefficient remains large and statistically significant after including all controls.

Because black offenders face more technical violations, they are also more likely to be revoked for breaking technical rules. The black effect for this outcome is roughly 10% of the non-black mean after including the full suite of control variables. However, the final two bars show that black offenders are also more likely to be arrested. These effects are correlated across geographies, as shown in Appendix Figure A3. In parts of the state where black offenders are more likely to be arrested relative to comparable non-black peers, they are also more likely to face technical violations. This pattern suggests that at least part of the racial disparities in technical violations may reflect the fact that likely reoffenders are also likely rule breakers.

### III MEASURING EFFECTIVENESS AND EQUITY

This section provides a framework for assessing the effectiveness and equity of revocation when viewed as a simple tool for predicting socially costly behavior. Similar ideas apply in other criminal justice contexts, including bail setting (Kleinberg et al., 2017a), parole release (Kuziemko, 2013), background screening. I then show how with the use of an instrument one can construct tests for racial differences in accuracy and type-I and type-II error rates, as well as a method for quantifying the contribution of any differences in error rates to aggregate disparities in outcomes.

#### IIIA *Static model*

Consider a simple one-period model. Let  $Y_i$  be a binary outcome that equals 1 if an offender is rearrested for a new criminal offense. Let  $R_i$  be an indicator for being revoked due to technical rule violations. Let potential criminal offending  $Y_i(0), Y_i(1)$  depend on whether or not an individual is revoked, with observed arrests  $Y_i = Y_i(R_i)$ . Throughout this section, I suppress an additional subscript  $s$  for probation spell, treating each person-spell observation as a separate unit to simplify exposition.<sup>15</sup>

My primary goal is to investigate racial differences in three key measures of the effectiveness of revocation. The first is predictive accuracy, or  $Pr(Y_i(0) = 1 | R_i = 1)$ . When accuracy is high, a large fraction of those revoked would otherwise reoffend. When it is close to the population mean  $Pr(Y_i(0) = 1)$ , then revocation has no signal value for reoffending. The second and third concepts provide alternative measures of effectiveness by reversing this conditional probability to examine type-I and type-II error rates, or  $Pr(R_i = 1 | Y_i(0) = 0)$  and  $Pr(R_i = 0 | Y_i(0) = 1)$ , respectively. Low error rates imply revocation is a better classifier of counterfactual reoffending risk.

Recent work on “algorithmic fairness” has explored how differences in accuracy and error rates correspond to conventional notions of bias or fairness (Corbett-Davies et al.,

14. Tables showing full regression results, including the effect of adding controls sequentially, are available starting with Appendix Table A2.

15. All results cluster standard errors by individual to account for potential within-person correlation in outcomes.

2017; Kleinberg et al., 2017a; Berk et al., 2018). A standard result in this literature is that it is impossible to simultaneously equalize type-I and type-II error rates and accuracy across groups unless an algorithm either perfectly predicts the outcome or outcome rates are the same across groups.<sup>16</sup> In what follows, I consider all three measures and demonstrate that race gaps in this setting do not reflect an edge case where two of the three measures are balanced.

How can accuracy and error rates be estimated? Extensions of results from the instrumental variables literature provide a solution. Suppose we have access to a binary instrument  $Z_i$ . Let potential revocation be indexed by the instrument as  $R_i(0), R_i(1)$ , and assume that the following standard 2SLS assumptions are satisfied (Angrist et al., 1996):

1. First stage:  $Pr(R_i = 1|Z_i = 1) < Pr(R_i = 1|Z_i = 0)$
2. Monotonicity:  $R_i(1) \leq R_i(0) \forall i$
3. Independence and exclusion:  $(Y_i(0), Y_i(1), R_i(0), R_i(1)) \perp\!\!\!\perp Z_i$

That is, the instrument weakly reduces the possibility of revocation for all individuals; is independent of potential reoffending and revocation; and affects reoffending only through whether or not an individual is revoked. Since serving supervised probation post-reform will be my instrument, this final assumption requires that the entire change in reoffending after the reform be attributable to changes in revocations. Imposing this exclusion restriction rules out offenders directly adjusting their reoffending behavior in response to the change in policy. Such responses are potentially plausible. For example, offenders might use more drugs when failed drug tests are punished less harshly, which could increase arrests. I provide tests supporting this assumption in Section IVC and, in the final part of the paper, I relax it and measure any behavioral responses directly.

Abadie (2002) shows that under assumptions 1-3 it is possible to characterize the mean reoffending rate of individuals shifted out of revocation due to the reform:

$$(1) \quad \frac{E[Y_i(1 - R_i)|Z_i = 1] - E[Y_i(1 - R_i)|Z_i = 0]}{E[1 - R_i|Z_i = 1] - E[1 - R_i|Z_i = 0]} = E[Y_i(0)|R_i(1) = 0, R_i(0) = 1]$$

This quantity measures the share of individuals revoked under the pre-reform rules ( $R_i(0) = 1$ ) but not post-reform rules ( $R_i(1) = 0$ ) with  $Y_i(0) = 1$ . It therefore characterizes the accuracy of the rules affected by the reform—drug and administrative violations—as tags for counterfactual reoffending. Probationers who break rules unaffected by the reform, namely absconding violations, continue to be revoked afterwards and hence have  $R_i(1) = R_i(0) = 1$ . The reform contains no information about accuracy for these “always takers.”

Because the numerator on the left-hand side of Equation (1) identifies  $Pr(Y_i(0) = 1, R_i(1) = 0, R_i(0) = 1)$ , changing the denominator and some minor manipulation makes it

16. To see this, note that accuracy is related to error rates as:

$$Pr(Y_i(0) = 1|R_i = 1) = \frac{1 - Pr(R_i = 0|Y_i(0) = 1)}{1 - Pr(R_i = 0|Y_i(0) = 1) + Pr(R_i = 1|Y_i(0) = 0) \frac{Pr(Y_i(0)=0)}{Pr(Y_i(0)=1)}}$$

Hence unless  $Pr(R_i = 1|Y_i(0) = 0)$  is zero for both groups or  $Pr(Y_i(0) = 1)$  is the same, accuracy will differ. In what follows I show that racial differences in revocation do not correspond to an edge case where two of these three measures are equalized.

also possible to estimate error rates for the population with  $R_i(1) = 0$ . For example, type-II error rates are:

$$Pr(R_i(0) = 0 \mid Y_i(0) = 1, R_i(1) = 0) = 1 - \frac{Pr(Y_i(0) = 1, R_i(1) = 0, R_i(0) = 1)}{Pr(Y_i(0) = 1, R_i(1) = 0)}$$

where  $Pr(Y_i(0) = 1, R_i(1) = 0)$  can be easily estimated from its sample analogue in the population with  $Z_i = 1$ . As with accuracy, error rates characterize the effectiveness of drug and administrative rules specifically. In Section V, I use a competing hazards model to estimate accuracy and error rates for all rules and the full population.

By estimating accuracy and error rates separately by race, one can easily compare these measures across groups. With race specific estimates of error rates, one can also decompose differences in technical revocation  $Pr(R_i = 1)$  into a share attributable to differences in error rates and a share attributable to differences in reoffending rates. Specifically, letting  $B_i \in \{0, 1\}$  denote race, we have:

$$(2) \quad \underbrace{Pr(R_i(0) = 1 \mid B_i = 1) - Pr(R_i(0) = 1 \mid B_i = 0)}_{\text{difference in technical revokes}} =$$

$$\sum_{k=0}^1 \underbrace{Pr(Y_i(0) = k \mid B_i = 0)}_{\text{non-black risk}} \underbrace{[Pr(R_i(0) = 1 \mid Y_i(0) = k, B_i = 1) - Pr(R_i(0) = 1 \mid Y_i(0) = k, B_i = 0)]}_{\text{difference in error / true positive rates}}$$

$$+ \underbrace{Pr(R_i(0) = 1 \mid Y_i(0) = k, B_i = 1)}_{\text{black error / true positive rates}} \underbrace{[Pr(Y_i(0) = k \mid B_i = 1) - Pr(Y_i(0) = k \mid B_i = 0)]}_{\text{difference in risk}}$$

Thus the total difference is comprised of a component driven by differences in risk ( $Pr(Y_i(0) = 1)$  and  $Pr(Y_i(0) = 0)$ ) and a component driven by differences in the likelihood offenders are revoked conditional on their potential reoffending status. As always with Oaxaca-style analyses, it is possible to construct alternative decompositions by adding and subtracting other composite terms (Oaxaca and Ransom, 1999). Here, I decompose the difference using the non-black risk distribution and the black error rates as the baseline. Results change little when doing the reverse.

### IIIB *Dynamic model*

The one-period model abstracts from the fact that probationers can be rearrested and revoked at any point in their spell. When implementing it, the choice of horizon over which outcomes are measured (e.g., the first year of a spell) may be consequential if racial differences in accuracy and error rates vary across horizons. Appendix Section A1 describes a simple extension to the one-period model that allows me to estimate accuracy and error rates at multiple horizons. To summarize the overall impact of differences in targeting vs. differences in reoffending risk, the decomposition in Equation (2) can also be extended to span many periods instead simply the binary indicator in the one-period model.

### IIIC *Connection to empirical setting*

As noted previously, some offenders no longer revoked due to the reform may have instead been confined for at most 90 days under a CRV, the new punishment introduced by the JRA legislation. Appendix Section A2 accounts for this feature of the reform and shows

that if CRVs are used exclusively as substitutes for revocations, the procedure described in Section IIIA estimates clearly interpretable accuracy and error rates.<sup>17</sup>  $Y_i(0)$ , however, reflects reoffending when subject to the alternative to revocation, namely a potential CRV, rather than no confinement whatsoever. Accuracy and error rates therefore refer to potential reoffending under this alternative policy. In what follows, however, I also consider several exercises that examine any direct effects of CRVs, including estimating effects after they were eliminated for misdemeanants in 2015.

The analysis that follows also uses a difference-in-differences strategy to account for any time trends in reoffending. Appendix Section A3 discusses the additional assumptions required to do so. As in any difference-in-differences analysis, these assumptions require outcomes in the control group to trend similarly to relevant populations in the treated group. In the empirical analyses that follow, however, I show that results change little when using a simple pre-post comparison of offenders starting their spells close to the reform rather than a differences-in-differences estimator.

## IV RESULTS

First, I analyze the effects of the 2011 JRA reform on revocation for technical violations and arrests over a one-year time horizon using a difference-in-differences estimator. This analysis implements the one-period model introduced in the previous subsection. This one-period analysis is also sufficient to conduct a simple cost-benefits analysis of the effectiveness of technical rules as tags for potential reoffenders and to compare the relative social return to revocation across race groups. I also describe several extensions and robustness checks.

### IVA *Unadjusted time series*

I analyze the 2011 JRA reform using two possible outcomes for each probation spell: 1) new criminal arrest; and 2) revocation for technical violation. These events are mutually exclusive—an offender cannot be technically revoked if they are arrested first by definition.<sup>18</sup> For each probationer, I measure which event occurs (if any) and the time to the event. I then calculate the shares of probationers revoked and arrested over the course of their spell. Figure II plots these shares in Panels A and B for three-month cohorts of supervised probationers. Each line represents the failure function for the cohort that started probation where the line intersects the x-axis. The line then tracks the share of this cohort experiencing the outcome over the first year of their spell. The leftmost line in Panel A, for example, plots the share of probationers starting their spells in the beginning of 2007 who were revoked over the next 365 days. By the end of that period, where the line ends, roughly 15% of the cohort was revoked for technical violations. Similar shares experience the same fate in each cohort for the next 12 quarters.

The reform’s effective date is marked with the black solid line. Cohorts beginning probation within a year of the reform begin to see reductions in revocation. These cohorts were affected because the reform’s limitations on technical imprisonment applied by the violation date and not the probationer’s start or offense date. Thus these cohorts spend

17. The appendix also discusses interviews conducted by the N.C. Sentencing Policy and Advisory Commission that support this assumption.

18. As noted earlier, technical revokes are defined as revocation for rule breaking with no intervening criminal arrest by regular NC law enforcement.

a portion of their spell under the new policy regime and see reductions in revocation as a result. The more time each cohorts spends under the new regime, the larger the reductions. Probationers who begin their spell after the reform are fully exposed to its changes. For these cohorts, revocation reduces to 9%, a 33% drop relative to the pre-reform mean.

The large decrease in revocation means many more probationers had the opportunity to be arrested instead. Panel B plots the share who did so. After a slight decline over several years, offending is relatively flat in the 4 quarters before the reform. It then jumps up slightly for spells interrupted by the reform and remains 1-2 p.p. higher afterwards. Thus while the reform sharply reduced revocation, these gains came at a cost.

This simple interrupted time series analysis may be misleading if selection into probation changed as a result of the reform. Appendix Figure A5 shows that the first threat is not a concern. Predicted offending rates formed using all available covariates are stable over the four years before and after the reform. Appendix Figure A6 shows that the quantity of offenders on supervised and unsupervised probation also did not change discretely around the reform, indicating that judges' sentencing behavior was unaffected.<sup>19</sup> Thus, although probation overall became more lenient after the reform, there is no evidence that either judges changed their sentencing behavior or potential offenders changed their crime choices in response.

## IVB *Difference-in-differences estimates*

It is also possible that time trends in reoffending would confound a simple pre-post analysis. To account for this, I use a difference-in-differences approach that compares supervised probationers' outcomes to unsupervised probationers'. Panel C of Figure II plots the difference in these groups' one-year revocation and arrest rates (i.e., the end-points of the lines in Panels A and B).<sup>20</sup> Specifically, it plots estimates of  $\beta_l^T$  from the linear regression:

$$(3) \quad Y_{is}^j = \alpha + \sum_{l \in \{-16, -5\} \cup \{-3, 16\}} 1\{S_{is} = l\}(\beta_l + \beta_l^T T_{is}) + e_i$$

where  $Y_{is}^j$  measures whether individual  $i$  in spell  $s$  experienced outcome  $j$  (either arrest or revocation),  $S_{is}$  measures how many quarters before or after the reform's effective date  $i$  started probation, and  $T_{is}$  is an indicator for being on supervised probation. The  $\beta_l^T$  effects are normalized relative to the cohort starting four quarters before the reform, the last group to spend the entirety of their first year of probation under the old regime.

Because unsupervised offenders are not assigned probation officers, less than 1% of them experience revocation in the first year of their spell. As a result, the reform had virtually no impact on this group. The blue line in Panel C thus closely tracks the declines in Panel A—decreases of roughly 6 p.p. after a prolonged period of no substantial changes. Because unsupervised probationers saw no decline in revocation, their arrest rates evolved smoothly over the reform. Beforehand, their outcomes tracked supervised probationers' closely for three plus years. The red line reflects this pattern, showing increases of 2 p.p. with no evidence of pre-trends.

To obtain point estimates of the reform's effects, I collapse Specification (3) to a simple difference-in-difference comparison using probation spells that begin 1-3 years before the

19. Appendix Table A9 shows that predicted reoffending rates and other core covariates are also not trending differentially in treated vs. control units.

20. The raw rates for unsupervised probationers are presented in Appendix Figure A7.

reform and 0-2 years afterwards, thus using two years of data before and after the reform while omitting cohorts whose first year was interrupted by the reform and were therefore only partially affected. These results are presented in Panel A of Table III. The estimated effect on revocation is 5.3 p.p and easily distinguishable from zero at conventional confidence levels. The increase in arrests is roughly 2 p.p. Viewed through the empirical model, these estimates imply that 30-40% of probationers spared revocation found themselves arrested instead. For both outcomes, it makes little difference whether demographic and criminal history controls are included.

Are these effects small or large? A simple benchmark for the reform’s expected effects uses the share of probationers arrested pre-reform, which was 29%. If a similar share of probationers spared revocation are arrested instead, we would expect offending to go up by roughly 1.6%. The observed increase falls slightly above this simple benchmark, suggesting individuals targeted by revocation are somewhat more risky than average. Since revocation occurs over the course of a probation spell, however, this benchmark is potentially too high. For example, in the extreme case where all revokes occur on day 355 of the spell, the reform would only give offenders *one* extra day to commit crimes in their first year, and finding any increase would be surprising. I return to this question in Section V, where I estimate arrest and revocation hazards directly and show that they are highly correlated across individuals.

The last two rows of Panel A present estimates of false positive and false negative rates. The estimated false negative rate shows that just 6.4% of potential reoffenders are caught by revocation due to the drug and administrative rules affected by the JRA reforms. The estimated false positive rate shows that 5.6% of non-offenders (over the one-year horizon), however, violate the same rules. Of course, many of these individuals may offend later, a fact I account for in the dynamic estimates discussed below. Nevertheless, in this simplified setting rules appear almost as likely to target non-reoffenders as reoffenders.

Remarkably, the reform’s impact on black offenders’ revocation was nearly twice as large as its impact on non-black offenders’. As a result, the reform eliminated raw racial disparities in revocation. Panel A of Figure III demonstrates this result by plotting revocation rates in the sample used for the difference-in-differences estimation separately by race. While black offenders were 30-40% more likely to face revocation over the first year of their spell before the reform, afterwards the race gap is reduced to less than 1%.

Because many more black offenders were spared revocation, one might expect arrests in the black population to increase more than in the non-black population after the reform. Panel B of Figure III shows that this did not happen. While more probationers in both groups were arrested after the reform, the racial gap does not change substantially. Race-specific difference-in-difference estimates in Panels B and C of Table III imply accuracy for non-black offenders is above 55%.<sup>21</sup> However, the corresponding figure for black offenders is only 31%. Estimates of false negative rates by race are similar—roughly 93%. But false positive rates are three times higher for black offenders, implying that far larger shares of black offenders who would not otherwise reoffend were revoked.<sup>22</sup>

These results are remarkably similar to those in Sakoda (2019)’s evaluation of a reform to post-release supervision in Kansas.<sup>23</sup> In 2000, Kansas eliminated post-release supervi-

21. Race-specific versions of Figure III Panel C are in Appendix Figure A8.

22. Criminal activity is measured using arrests. If black offenders are also policed more aggressively while on probation, disparities in accuracy and error rates for arrests may *understate* disparities in accuracy for total unobserved criminal offending.

23. Post-release supervision is closely related to parole. Parolees typically serve the remainder of their original incarceration sentence under community supervision after being released from prison. Post-release



sion for a subset of offenders incarcerated for a prior probation violation. Using a similar difference-in-differences design, Sakoda (2019) finds that the reform decreased three-year reincarceration rates by 31 p.p. and eliminated race gaps in reincarceration, but had no impact on new felony convictions that resulted in prison time. The concordant results suggest the racially disparate impacts of revocation found here are not a special feature of North Carolina’s probation regime.

Table IV uses these results to conduct the simple Oaxaca decomposition exercise described previously. This analysis measures the relative contributions of risk (i.e.,  $Pr(Y_i(0) = 1)$ ) and targeting (i.e.,  $Pr(R_i = 1|Y_i(0) = 1)$ ) to aggregate racial gaps in revocation.<sup>24</sup> As expected, the first two rows show that rates of revocation and offending are both higher in the black population. The next two rows, however, show that while black offenders’ higher likelihood of reoffending contributes slightly, it is more than fully offset by harsh treatment of non-offenders. The bulk of differences in revocation are driven by differences in how often non-offenders are revoked, which explains 105% of the aggregate gap.

The appendix contains several robustness exercises and extensions of these main results. Appendix Table A10 drops the control group and estimates single-difference effects. Estimated disparities change little, with accuracy for black offenders roughly half that of non-black offenders and type-I error rates three times larger. Appendix Table A16 tests for sensitivity to the data window around the reform used in estimation. Disparities remain similar as the window is lengthened, although accuracy estimates increase somewhat for both race groups.<sup>25</sup> Finally, Appendix Table A14 shows that the increase in crimes by type do not differ substantially across the two race groups. In fact, the absolute increase in felony offenses is *smaller* in the black population than in the non-black population. It therefore does not appear that black probationers targeted by technical violations pose lower average risk, but higher risk for more socially costly crimes such as felonies.

## IVC Testing for behavioral effects

Interpreting the effects of the reform through the empirical model presented earlier requires that offenders do not respond directly to the change in policy. Behavioral responses, however, find little support in the data. Table V demonstrates this by estimating a post-reform effect in Cox proportional hazards models for arrests and rule breaking. When studying arrests in Columns 1 and 2, these regressions treat any technical rule violation as a source of censoring. Doing so removes any arrests that occur after a rule violation and hence may have been censored by revocation pre-reform. If no increases in arrest hazards are detected, this implies that increases in offending post-reform are explained by the mechanical change in revocation rather than offenders being rearrested more frequently or earlier in their spells.<sup>26</sup>

Analogous regressions can be estimated to test whether rule violations themselves increase after the reform. The results show no change in any violation rates (Columns 3 and 4), fees and fines violations (Columns 5 and 6), or drug violations specifically (Columns 7 and 8). While perhaps surprising, these results are consistent with a series of randomized

supervision is a fixed period of community supervision that offenders serve after finishing their active prison sentence.

24. Appendix Section A4 provides complete details on how the decomposition is calculated.

25. Similar results for offenders who begin probation shortly after the reform and thus largely committed offenses before December 1, 2011 are also reassuring that JRA’s other changes to probation, which applied by original offense date, do not explain the estimated racial disparities.

26. See Appendix Figure A4 for a graphical illustration.

controlled trials demonstrating that probationers’ offending and drug test failure rates do not respond to stricter monitoring or more intensive probation conditions (Hennigan et al., 2010; Barnes et al., 2012; Boyle et al., 2013; Hyatt and Barnes, 2017). Since the reform also provided the option for short periods of confinement in place of full revocation for technical violations, offenders may have also viewed the potential reduction in punishments induced by the reform as uncertain or limited.

Nevertheless, it remains possible that exclusion does not hold exactly. Appendix Section A5 examines the sensitivity of the main results to violations in the spirit of Conley et al. (2012) and Rambachan and Roth (2020). Racial differences in accuracy persist so long as violations of exclusion are not large and opposite signed for each race group.

### ***IVD Impacts of CRVs***

As noted earlier, JRA introduced the option to impose short periods of confinement (CRVs) as a substitute for revocation. Reoffending rates captured in the preceding analysis therefore reflect outcomes under this alternative policy rather than no confinement at all. Racial differences in accuracy and error rates also reflect outcomes under this alternative, as discussed in Appendix A2. It is possible that results would be different if there were no option to impose CRVs.

Several exercises, however, suggest that CRVs do not explain the estimated racial disparities. First, for felony offenders, it is possible to estimate the effect of the reform on total days incarcerated for technical violations through either revocation or a CRV. Appendix Table A11 shows that the reduced-form effect on days incarcerated for black offenders was roughly twice as large as for non-black offenders and that black and non-black offenders no longer revoked in their first year due to the reform experienced 203 and 172 fewer days of technical incarceration, respectively. The larger decreases in incarceration for black offenders make their smaller observed increases in offending even more surprising.<sup>27</sup> Second, CRVs were eliminated for misdemeanor offenders in 2015. Appendix Table A12 shows that comparing misdemeanor offenders on probation before the reform to those serving in 2016 reveals the same patterns of racially disparate impacts. Finally, Appendix Table A13 extends the horizon over which reoffending is measured by 90 days to account for any potential incapacitation due to CRVs. The results again show the same pattern of effects.

### ***IVE Impacts of specific rule types***

Which types of technical rules generate these racial differences? This question is difficult to answer without additional assumptions. The reform impacted a bundle of technical rules. Although decreases in revocation for each rule type are observable, only overall increases in arrests can be estimated. However, it is always possible to examine the correlation between arrests and rule violations among offenders not revoked. Appendix Table A18 does so by reporting rule violation rates for probationers arrested, not arrested, and revoked in the post-reform data. The estimates show that black probationers who do *not* reoffend are more likely to break all rule types than black probationers who do (i.e., false positive > true positive rates). The differences are particularly stark for fees and fines violations: 36%

27. One could also define treatment as any technical incarceration (due to a CRV or revoke). Doing so shows similar patterns of racial disparities, but suffers from a clear exclusion restriction issue because some offenders are shifted from revoke to CRV (hence no change in this definition of treatment) but still experience large decreases in total time incarcerated.

of black probationers who do not reoffend incur a financial violation, vs. 13% of black probationers who do. Unpacking the full population relationship between arrests and rule violations, however, requires accounting for the potential reoffending of individuals revoked for rule breaking. Section VD uses a competing hazard model to do so.

### IVF Triple-difference estimates

The previous results demonstrate that revocation has remarkably different impacts on black and non-black offenders. However, black and non-black offenders may differ in important observable characteristics, including their age and gender composition, extent of criminal history, and geographic distribution throughout North Carolina. To examine how sensitive the previous results are to accounting for such observable differences, I estimate a triple-difference version of specification (3):

$$(4) \quad Y_{is}^j = \underbrace{\alpha + \beta_1 T_{is} + \beta_2 P_{is} + \beta_3 T_{is} P_{is}}_{\text{D-in-D regressors}} + \underbrace{B_i(\beta_4 \alpha + \beta_5 T_{is} + \beta_6 P_{is} + \beta_7 T_{is} P_{is})}_{\text{Interaction with black indicator}} + \underbrace{X_{is}(\beta_8 \alpha + \beta_9 T_{is} + \beta_{10} P_{is} + \beta_{11} T_{is} P_{is})}_{\text{Adjustments for observables}} + e_i$$

where  $P_{is} = 1\{S_{is} \geq 0\}$ , i.e., a “post” indicator,  $B_i = 1$  if offender  $i$  is black, and  $X_{is}$  is a set of observable characteristics that does not include race.  $\beta_7$  captures differential changes in the outcome  $Y_{is}^j$  for treated black vs. non-black offenders before vs. after the reform relative to changes experienced before vs. after the reform by untreated offenders. If  $\beta_7 = 0$ , then “post-x-treat” coefficients in a standard difference-in-differences specification estimated separately for black and non-black offenders would be identical. Including  $X_{is}$  allows me to make this comparison after adjusting for differences in observable characteristics between black and non-black offenders.

Table VI reports estimates of  $\beta_7$ , labeled “treat-x-post-x-black”, and  $\beta_3$ , labeled “treat-x-post” for varying sets of controls  $X_{is}$ . The first two columns omit  $X_{is}$  entirely. As shown earlier, black offenders experience much larger declines in incarceration for rule breaking but see increases in reoffending that are indistinguishable from non-black offenders’.<sup>28</sup> Columns 3 and 4 add demographic controls, so that only black and non-black offenders of the same age and gender are compared. Black offenders continue to see roughly two times larger decreases in incarceration, but identical increases in reoffending. The next sets of column pairs add criminal history controls, indicators for the probation district where the offender is being supervised, and indicators for zip code of residence at the time of the original conviction. After adjusting for these factors, black offenders continue to see substantially larger decreases in incarceration but no different changes in reoffending.

These results need not imply that *race itself* drives the differential impact of probation’s technical rules. As argued in Section IVI below, the evidence in fact suggests that racial disparities in this setting do not arise due to racial bias on the part of police, judges, or probation officers, and instead reflect differences in behavior between black and non-black offenders. However, Table VI shows that such differences are not easily explained with observable characteristics, including reasonable proxies for income such as residential

28. The post-x-treat coefficients reported here are identical to the post-x-treat estimates in Panel B of Table III columns 1 and 3. Adding the treat-x-post-x-black coefficients reproduces the post-x-treat estimates in Panel C columns 1 and 3.

neighborhood. This suggests that the behavioral differences between black and non-black offenders that drive technical revocations' disparate impact may reflect other more nuanced and contextual factors, such as access to informal credit that could be used to pay off fees and fines. The stability of effects to covariates also suggest that results are not a special feature of the particular black and non-black populations sentenced to probation under the current criminal justice system.

### IVG Cost-benefit analysis

When the state revokes an offender for technical violations, it pays on average \$100 a day to do so.<sup>29</sup> If the state instead opts to leave the offender in the community, she may then be arrested and be sentenced to incarceration as a result. The social value of technically revoking individual  $i$  can thus be written as:

$$(5) \quad V_i = \underbrace{-J_i}_{\text{Cost of tech. incar.}} + \underbrace{\Pr(Y_i = 1 | R_i = 0)}_{\text{Pr(offend) if not incar.}} \left[ \underbrace{E[U(Y_i) | R_i = 0, Y_i = 1]}_{\text{Cost of crime}} + \underbrace{J'_i}_{\text{Cost of new sent.}} \right]$$

where  $J_i$  is the cost of the revocation,  $R_i$  and  $Y_i$ , as before, are indicators for revocation and reoffending,  $U(Y_i)$  represents the social cost of this crime, and  $J'_i$  represents the total cost of incarceration as a result of the new arrest, including any resulting revocation.

Revoking technical violations for a group offenders is beneficial if  $E[V_i] > 0$ . I assess this criterion for offenders affected by the 2011 JRA reforms in two ways. First, I use changes in observed costs of revocation and offending rates over a fixed horizon to back out a “break-even”  $E[U(Y_i) | R_i = 0, Y_i = 1]$  that sets  $E[V_i] = 0$  for this population. That is, I solve for:

$$(6) \quad E[U(Y_i) | R_i = 0, Y_i = 1] = \frac{\Delta E[-J_i \cdot R_i] - \Delta E[(1 - R_i)J'_i]}{\Delta E[Y_i]}$$

This exercise asks what is the *minimum* social cost of arrests needed to justify the use of revocation for the drug and administrative rules impacted by the reform. The numerator captures the change in net incarceration costs—spending on revocation minus spending on arrest-driven incarceration. The denominator divides this gap by the increase in reoffending to arrive at break-even valuation for these marginal offenses. I consider costs and benefits of revocation that begins and arrests that occur in the first year of a probation spell. To measure incarceration costs, I use the length of suspended sentences and sentences for new criminal activity recorded in AOC data.<sup>30</sup>

Costs of incarceration are assessed using the average costs of per prisoner-day. For small shifts in incarceration rates, marginal costs may be more appropriate. Lower costs of incarceration serve to scale *down* break-even costs of arrests. If the true marginal cost is \$50 per day, for example, break-even valuations would be half as large as those presented here. In this context, true costs may be close to average costs, however. The JRA reform lead to substantial declines in incarceration that enabled the DPS to close several facilities

29. 2018 average daily cost per inmate for the NC DPS (<https://www.ncdps.gov/adult-corrections/cost-of-corrections>). Supervision costs roughly \$5 a day in 2018.

30. This allows me to capture the costs of incarceration in local jails for misdemeanants, which is not recorded in the state prison incarceration data from DPS.

(Hall et al., 2015), although part of these decreases were driven by other changes beyond probation.

In a second approach, I use existing estimates from the literature to benchmark crime costs and compare it to these break-even values. This analysis assigns a cost to each category of arrest ranging from \$500 (for simple drug possession) to close to \$20 million (for homicides) primarily sourced from Cohen et al. (2011).<sup>31</sup> I then compare the change in net incarceration costs due to the reform to estimated increases in costs of crime.

This analysis omits several other factors that might contribute to the aggregate costs and benefits of revocation. In particular, the foregone earnings of incarcerated offenders, the utility costs of imprisonment, and the court costs associated with processing rule-driven incarceration are excluded. The excluded potential benefits mainly relate to deterrence effects. As shown earlier, however, there is little evidence that the reform impacted the perceived punitiveness of probation enough to shift probationers' offending calculus. Nor is there any change in technical rule compliance rates after the reform, including for payment of fees or fines.<sup>32</sup> On net, therefore, I view this analysis as providing a lower bound on costs while capturing most potential benefits.

Importantly, these cost-benefit calculations also place no weight on racial equity. Since the reform dramatically reduced racial gaps in revocation, this is a potentially important factor. Indeed, the more policy makers value reducing racial disparities, the more attractive the reform becomes regardless of its impact on arrests. A full social welfare analysis of the reform—including putting a price on racial equity—is beyond the scope of this paper, however.

The results are reported in Table VII. The first column reports the change in spending on revocation activated in the first year of a probation spell after the reform took effect. This declined by \$633 per probationer on average. The second column reports the increase in costs of incarceration attributable to new arrests in the first year of a spell. These increases are relatively small because the majority of new arrests induced by the reform do not merit an actual prison sentence. The estimates thus imply that for every dollar the state spent on revocation, it saved roughly 35 cents it would have spent on prison costs anyway.

Column 3 reports the implied break-even valuations discussed above. These average about \$40k per offense. Although this may seem relatively low, consider that the modal offense committed by a probationer is a relatively minor misdemeanor. In fact, excluding all misdemeanor and traffic offenses raises the marginal valuation to \$100k (Column 4). Columns 5 and 6 report the estimated costs of new crimes generated by the reform. Unfortunately, due to the wide dispersion in reported costs of crime, these estimates are relatively noisy. The point estimates, however, suggest that costs may fall at or below break-even valuations.

The remainder of Table VII repeats the same exercise for various sub-populations. The second and third rows provide a concise summary of the degree to which drug and administrative revocations target black offenders more aggressively. The decrease in spending on revocation in the black population is roughly twice as large, while increases in the costs of incarceration attributable to new arrests are only slightly larger. Combined with simi-

31. See Appendix Table A26 for a detailed list of crime costs and their sources. Each arrest is assigned a lower and upper bound for costs based on existing estimates and the categorization of the offense.

32. Due to the decline in revocation, however, offenders spend more time on probation post-reform and have more opportunities to accumulate technical violations. Appendix Table A19 shows that on average offenders accumulated 0.05 more violations in their first year of probation due to the reform. This analysis does not account for any social cost imposed by these violations.

lar increases in reoffending rates for both groups shown earlier, the result is that implied break-even valuations for black offenders are 2-3 times larger than for non-black offenders. Unfortunately, estimates in Columns 5 and 6 are too noisy to ask whether differences in costs of crime justify these disparities. However, racial gaps in break-even valuations are even larger when only felony offenses are considered in Column 4, suggesting that differences in the severity of crime committed are unlikely to justify the gap. The final two rows of Table VII shows that similar but more extreme patterns hold when considering black and non-black men.

#### ***IVH Dynamics model estimates***

The analysis thus far has treated the first year of probation as a single period. In the appendix, I generalize these results using the dynamic model described in Section IIIB. Doing so is potentially important if there are large racial differences in the timing of reoffending, or if racial gaps at one year are counterbalanced by differences at other horizons. Appendix Figure A11 plots type-II error rates for black and non-black offenders over a three-year horizon.<sup>33</sup> Lower values for black offenders indicate that rules target a smaller fraction of black potential reoffenders at each time  $k$ . The final point for each group reflects the share of probationers who would not reoffend within three years of starting probation but were revoked for technical violations, or type-I error rates at a three-year horizon. Type-I error is significantly higher for black offenders.

Appendix Table A17 summarizes the impact of differences in error rates for aggregate disparities in revocations using the decomposition exercise presented previously for the one-period model. The results show that black offenders do reoffend more often and earlier in their spells. As a result, differences in risk explain a small portion of race gaps in technical revocation. However, differences in the likelihood of revocation conditional on reoffending explain the majority of the aggregate disparity. Black revocation for drug and administrative violations would have been virtually identical to non-black revocation if they had similar conditional likelihoods of revocation as non-black offenders but their reoffending distribution were left the same.

#### ***IVI Behaviors or biased responses?***

In general, racial disparities in technical violations could arise for two reasons. First, black offenders may be more likely to exhibit the proscribed behaviors. For example, black offenders may have more limited wealth and income and thus find it more difficult to pay fees and fines. Likewise, some populations may have less access to transport, making it more difficult to report to probation officers. In either case, disparities reflect genuine differences in behavior across the populations, whatever their root cause. Alternatively, probation officers and judges may respond more aggressively to identical behaviors when a probationer is black.

Several pieces of evidence suggest that differences in behaviors rather than responses to them are important for explaining the observed disparities. First, there is limited cross-officer variation in black offenders' likelihood of technical violations relative to non-blacks'. As shown in Appendix Table A8, controlling for assigned officer has no measurable impact on

33. I construct estimates of  $k$ -specific accuracy and error rates binning periods into 90 day intervals to gain precision. I thus test for disparities conditioning on reoffending falling somewhere within this interval rather than at  $k$  exactly, although results are not sensitive to the exact bin size.

the black effect for technical violations. Relatedly, Appendix Table A8 also shows that there is no consistent evidence of same-race effects, a pattern common in other criminal justice contexts where decision makers exercise wide discretion (e.g., West (2018)). Black officers are as likely to cite black offenders for administrative violations as non-black offenders.<sup>34</sup> It remains possible, of course, that officers are uniformly biased against black offenders, which would not be detected in these across-officer comparisons.

Second, racial disparities are large for technical violation categories where officers have relatively limited discretion as well as those where they have more. For example, relative to the violation’s mean frequency, black offenders are equally more likely to face violations for not reporting as for failing drug tests. While officers could fairly easily ignore a forgotten meeting, drug tests are initiated with an automated form produced by the DPS’s offender tracking computer system and thus harder to sweep under the rug. Black effects divided by the non-black mean for all violation categories are presented in Appendix Figure A2. These patterns are consistent with officers closely following detailed guidelines in the NC Department of Community Corrections’ policy manual, which specify appropriate responses to different probationer behaviors.

Third, racial disparities in revocation for technical violations are entirely driven by how often offenders pick up violations, not how those violations are punished. Conditional on the violation type, probation officers are equally likely to recommend revocation for black and non-black offenders and judges are equally likely to grant it, as shown in Appendix Table A7. In fact, simple fixed effects capturing violation types explains 40% of the variation in revocations, implying limited discretion overall in revocation punishments for technical violations.

## V COMPETING HAZARDS ANALYSIS

The previous results demonstrate that revocation for drug and administrative violations targets black offenders substantially more aggressively. This analysis, however, leaves several important questions unanswered, including the importance of arrest and violation timing, the size of any behavioral responses, and the effects of specific rule types. Answering these questions is difficult using the quasi-experimental variation alone. Estimating timing, for example, requires separating the impacts of state dependence and unobserved heterogeneity on observed variation in behaviors over the course of a spell. Because the reform impacted a bundle of technical rules simultaneously, it is also difficult to estimate the effects of specific rule types directly. This section therefore introduces a semi-parametric model of competing hazards for violations, revocations, and reoffending that provides a solution by explicitly accounting for both observed and unobserved heterogeneity in all behaviors.

### VA Setup

I model individuals’ latent hazards of new criminal arrest,  $Y_{is}^*$ , and revocation for technical rule breaking,  $R_{is}^*$ , using a mixed logit specification. Specifically, the discrete-time

34. For drug violations, black officers treat black offenders slightly *more* harshly on average.



hazards for individual  $i$  in period  $t$  of their  $s$ th probation spell are given by:

$$(7) \quad Pr(Y_{is}^* = t | Y_{is}^* \geq t, X_{is}, U_i^Y) = \Lambda(\theta_0^Y(t) + X'_{ist}\beta^Y + U_i^Y)$$

$$(8) \quad Pr(R_{is}^* = t | R_{is}^* \geq t, X_{is}, U_i^R) = \Lambda(\theta_0^R(t) + X'_{ist}\beta^R + U_i^R)$$

$\theta_0^Y(t)$  and  $\theta_0^R(t)$  are unrestricted baseline hazards for each outcome shared by all individuals.<sup>35</sup>  $X_{ist}$  are individual covariates, such as age and criminal history, that potentially vary between and within spells.  $U_i^Y$  and  $U_i^R$  are unobserved, individual-specific heterogeneity terms that will be treated as random effects. Both are constant across spells, an assumption that provides an important source of identification discussed further below. However, because  $X_{ist}$  can include covariates such as the number of previous spells, age, or calendar time, the same individual need not have the same hazard in repeated spells. In essence, therefore, only relative risk across individuals with the same observables is assumed constant across repeat spells.

This model can be viewed as a logit version of the canonical proportional hazard model introduced by Cox (1972).<sup>36</sup> The two outcomes' hazards can be correlated through observables through  $\beta^Y$  and  $\beta^R$  and unobserved heterogeneity through  $U_i^R$  and  $U_i^Y$ . With knowledge of  $\theta_0^Y, \theta_0^R, \beta^Y, \beta^R$  and the joint distribution of  $U_i^Y$  and  $U_i^R$ , it is straightforward to characterize how the risk of criminal arrest and technical revocation are related. By including in  $X_{ist}$  an indicator for whether period  $t$  falls after the 2011 reform took effect, one can also measure any behavioral responses in reoffending to the change in policy and the large decrease in the risk of revocation it induced.

Identification of  $\theta_0^Y, \theta_0^R, \beta^Y$ , and  $\beta^R$  comes from the empirical hazards. Identification of the unobserved heterogeneity  $U_i^Y$  and  $U_i^R$  comes from the joint distribution of survival times across multiple spells.<sup>37</sup> If there is no unobserved heterogeneity, then the joint distribution should factor into the product of marginals for each spell. If, on the other hand, individuals who are arrested quickly in their first spell are also likely to be arrested quickly in their second, there must be a sub-population with high  $U_i^Y$ . The same logic applies to the joint distribution of survival times across arrests and technical revocations. Behavioral responses, on the other hand, are identified by the impacts of the reform. If reoffending increases by *more* than what would be predicted by the decrease in revocation alone, then some behavioral response to the reform is necessary to rationalize the data.<sup>38</sup>

## ***VB Estimation***

I estimate the model separately by race (black vs. non-black) and gender (male vs. female). Doing so allows the joint distribution of unobserved heterogeneity, as well as the impact of observable characteristics, to have unrestricted differences across these groups. To capture baseline hazards, I include a fifth-order polynomial in duration. Rather than incorporating untreated probationers to account for time variation in offending, I include

35. In practice, I estimate a high degree polynomial in duration, although results are similar if indicators for fixed intervals are used instead.

36. In this case, the log odds of arrest in period  $t$  conditional on not being arrested before  $t$  are linear in the baseline hazard, covariates, and unobserved heterogeneity (and likewise for revocation). Efron (1988) studies a logit version of discrete time hazard models.

37. Individuals have multiple spells because they frequently reoffend and are re-sentenced to probation. As shown in Table I, there are 1.33 spells per person in the treated sample.

38. Appendix Section A6 provides additional details on identification.

simple time trends in the intercept, although results are not sensitive to this choice. Observables  $X_{ist}$  include indicators for whether the individual has multiple spells, a spell indicator interacted with duration (allowing the baseline hazard to differ in the first vs. second spell), a third-order polynomial in age, and an indicator for whether period  $t$  falls after the reform. I discretize time to the weekly level for computational speed, censor spells after three years, and use all spells starting up to three years after the JRA reform.

To model the unobserved heterogeneity, I follow Heckman and Singer (1984) and approximate the joint distribution of  $U_i^Y$  and  $U_i^R$  with mass points. That is, each individual belongs to one of  $K$  types, each with different values of  $U_k^Y$  and  $U_k^R$ . I then estimate the population shares of each type and their values of  $U_k^Y$  and  $U_k^R$ . While I normalize types so that the first has the lowest unobserved reoffending risk, I make no restrictions on the relative risk of rule violations across types. This allows, for example, types with very high offending risk to have either high or low risk of technical rule breaking. In the baseline estimates, I allow for four total types.<sup>39</sup>

The likelihood in finite mixture models is not concave, making maximization difficult. To ensure the results reflect a global optimum, I estimate the model many times using a large number of random starting points and keep the results that produce the largest value of the log likelihood.<sup>40</sup> To ensure that the results are robust to sensible alternative choices, I also estimate a version of the model with continuous heterogeneity that is a generalization of a standard bivariate probit. This version specifies that:

$$\begin{pmatrix} U_i^Y \\ U_i^R \end{pmatrix} \sim N(\alpha, \Sigma)$$

The continuous heterogeneity version has the convenient feature that unobserved racial differences in the correlation between arrest and rule-breaking risks are neatly summarized by the covariance terms in  $\Sigma$ .

## VC Results

The results support the conclusions from the previous analysis and reveal several important new insights. Figure IV, for example, plots average outcome-specific hazards for black and non-black men over the first three years of a spell. As expected, black men have both higher arrest and technical revocation hazards. The degree of duration dependence in arrest hazards for both groups is relatively minor, decreasing roughly 0.3 percentage points over the first year before flattening out slightly. The risk of technical revocation, however, peaks mid-way through the first year of a spell before declining to close to zero.

The estimated distributions of  $U_i^R$  and  $U_i^Y$  show that unobserved heterogeneity is an important feature of the data. Among black men, for example, the lowest reoffending risk type comprises 11% of the population and has a 3.5 log point lower weekly odds of offending than the highest risk type, which makes up 8% of the population.<sup>41</sup> Unobserved arrest risk has a strong correlation with unobserved revocation risk. Black men with the highest reoffending risk, for example, have 0.92 log point higher weekly odds of technical revocation than those with the lowest reoffending risk. Low-risk non-black men have even lower risk

39. Adding additional types increases the likelihood but does not change any of the core conclusions discussed below. As the number of types increases, however, optimization becomes more likely to become stuck in local maxima. Results with up to six types are available upon request.

40. Additional details are provided in Appendix Section A6.

41. Full parameter estimates are reserved for Appendix Tables A20 and A21.

of technical revocation, with 7% of the population belonging to a type that is relatively low arrest risk and virtually never subject to technical revocation.

This combination of state dependence and unobserved heterogeneity helps explain why technical rules are not more useful tools for identifying potential reoffenders and produced large error rates in the reduced-form analysis. The highest risk probationers are significantly more likely to reoffend early in their spells. Over time, the population that remains on probation shifts towards individuals with lower risk of reoffending. Thus, when the risk of technical revocation peaks, the riskiest offenders have already “selected out” of the pool still on probation.

The model also agrees that black offenders are targeted more aggressively by technical rules. Unlike in the previous reduced-form analysis, these estimates capture the effects of all technical violations and not just those impacted by the 2011 reform. Appendix Figure A12 demonstrates this by plotting  $Pr(R_{is}^* < Y_{is}^* | Y_{is}^* = k)$  separately for black and non-black men. Regardless of when individuals would otherwise be re-arrested, black men are substantially more likely to be subject to technical revocation. Black men who would not be rearrested within three years of starting probation (and possibly would never be) are nearly twice as likely to be revoked than similar non-black men.<sup>42</sup>

Estimates of the effect of the reform on hazards show large impacts on the odds of revocation, which is 0.50 and 0.39 log points lower for black and non-black men, respectively, after the change in the law. Consistent with the assumptions in the reduced form analysis, however, the reform had limited impacts on the underlying propensity to reoffend. Estimates for both genders are small and positive. Appendix Figure A16 plots the implied effect of these responses on average hazards. Pre- and post-reform arrest hazards are barely distinguishable; the mean difference is less than 0.1 p.p. at the pre-reform distribution of covariates. Moreover, these responses diminish as more flexible controls for calendar time are included in the model or more types are added.

Appendix Section A6 contains several additional results and validation tests of the model. These include a comparison of the model’s baseline hazards to Kaplan and Meier (1958) estimates of the same objects, a test of whether the model can accurately predict the impacts of the 2011 reform, and estimates of the model using continuous heterogeneity rather than types. These results support the conclusions of the baseline specification.

## VD *Disaggregating violation types*

To account for multiple types of rules, one could simply break up  $R_{is}^*$  into separate hazards for revocation for breaking drug-related rules, absconding, etc., turning the two-outcome competing risk model into an N-outcome model. Doing so, however, would throw out useful information. Because not all rule breaking results in revocation, offenders often break a rule, are punished with a warning, and are rearrested later in their spell. If this happens more often for offenders who break drug rules than for offenders who fail to pay fees and fines, then the former behavior may be more strongly connected to reoffending risk than the later.

To make use of this variation, I decompose the latent risk of technical revocation into

42. Part of this racial difference is driven by differences in observed characteristics, such as age and criminal history, while the remainder is driven by unobserved heterogeneity. Appendix Figure A17 shows that unobserved heterogeneity is responsible for the majority of racial differences. This plot holds each race group’s covariates fixed at the sample mean. The patterns change little.

two components:

$$(9) \quad Pr(R_{is}^* = t | R_{is}^* \geq t) = Pr(V_{ist}^k = 1 | R_{is}^* \geq t) Pr(I_{ist} = 1 | V_{ist}^k = 1, R_{is}^* \geq t)$$

Here,  $V_{ist}^k = 1$  is an indicator for breaking a technical rule of type  $k$  at duration  $t$  in spell  $s$  and  $I_{ist}$  is an indicator for being revoked as a result. An individual can have  $V_{ist}^k = 1$  multiple times within a spell, or have  $V_{ist}^k = 1$  and be rearrested subsequently. I model both components using a similar logit structure:

$$(10) \quad Pr(V_{ist}^k = 1 | X_{ist}, U_i^{V^k}, R_{is}^* \geq t) = \Lambda \left( \theta_0^{V^k}(t) + X'_{ist} \beta^{V^k} + U_i^{V^k} \right)$$

$$(11) \quad Pr(I_{ist} = 1 | V_{ist}^k = 1, X_{ist}, U_i^{V^k}, R_{is}^* \geq t) = \Lambda \left( \theta_0^I(t) + X'_{ist} \beta^I \right)$$

The  $\theta_0$  terms describe how the risk of type  $k$  rule violations and revocation evolves within a spell. The relationship between  $\beta^Y$  and  $\beta^{V^k}$  determines how *observable* characteristics drive correlations between the risk of breaking type  $k$  rules and the risk of reoffending. The relationship between  $U_i^{V^k}$  and  $U_i^Y$  determines *unobservable* correlations in the risk of reoffending and rule-breaking.<sup>43</sup> I continue to approximate the distribution of unobserved heterogeneity components using mass points. Since there are four types of violations (along with the possibility of arrest) each type now has five separate  $U_i$  components. I also include the same covariates as before, but allow the violation type and the number of previous violations to affect the risk of revocation in Equation (11).

I break rule violations into four types: reporting violations, such as absconding and missing regular meetings with a probation officer; drug and alcohol violations, such as failing a drug screen; fees and fines violations; and all others.<sup>44</sup> While these categories cover different behaviors, some rule violations mechanically produce others. For example, 81% offenders who stop reporting are also cited for unpaid fees, since it is difficult to pay fees if one has fled supervision. Offenders who fail a drug test are billed for the costs of the test, leading to more unpaid fees. To better capture the root behavior that lead to the violation, I code violations as reporting violations if there is any reporting violation, as drug violations if there is a drug violation but no reporting violation, and as fees and fines violations if there is a fee and fine violation but no drug or reporting violations. Results change little, however, if I do not recode violations to make them mutually exclusive and simply use all reporting, drug, and fees and fines violations in the data.

Parameter estimates from this version of the model for men are reserved for Appendix Tables A24 and A25. These estimates show substantial evidence of unobserved heterogeneity and state dependence. The estimates also show economically small increases in the risk of reoffending as a result of the change in policy and small changes in the risk of rule violations. Fees and fines violations, for example, show small and statistically insignificant *declines* in frequency after to the reform. Revocation risk conditional on breaking a rule, however, drops dramatically. This extension of the model therefore also supports the assumptions made earlier that the reform primarily impacts incarceration risk conditional on breaking a rule, but not offenders' reoffending or rule-breaking behavior.

43. Consistent with the reduced-form results showing that the decision to incarcerate conditional on breaking a rule is largely formulaic, unobservables do not enter the likelihood of punishment for rule breaking, although the model could easily be extended to allow this.

44. Other rules include violations rarely charged, such as failing to pursue vocational training or contacting a victim.

To study how each individual violation type relates to reoffending risk, I simulate the effects of enforcing particular subsets of rule types (e.g., just drug violations, drugs and fees and fines, etc.) with revocation. Figure V shows the results of this exercise. The x-axis plots the share of probationers who would reoffend over the first three years of a spell but break the enforced subset of technical rules beforehand (the true positive rate for the regime). The y-axis plots the share of non-reoffenders over the same period who do not violate any rules (the true negative rate). The technical rule “regime” enforced in each point is indicated in the labels: “F” for fees / fines, “D” for drug / alcohol, “R” for reporting, and “O” for all other rules. The regime’s effectiveness improves moving to the top-right corner of the graph, indicating that the rules catch more would-be reoffenders and imprison fewer non-reoffenders. The dotted gray line starts at (0,1) and has a slope of -1. This line reflects what would be achieved by randomly revoking a fraction of probationers at the start of their spells, which naturally would catch equal shares of reoffenders and non-reoffenders.

Figure V illustrates several interesting features of technical rules. First, most regimes for black men are interior to those of non-black men, indicating that all rule types generally have a tougher time discriminating between black offenders and non-reoffenders. Second, using rules related to fees and fines is almost always dominated by not doing so, and particularly for black offenders. Switching from enforcing fees to enforcing drug violations, for example, would result in catching 2-3 p.p. more would-be black reoffenders and imprisoning 12 p.p. fewer black non-reoffenders. Hence, reducing revocation for fees and fines not only improves effectiveness but also reduces disparities, as in North Carolina’s 2011 reform. Indeed, the post-reform regime for black men (“R”) now does better than random chance. Third, drug and reporting rules both appear to perform similarly. Using them in combination tends to simply increase the aggressiveness of the regime overall, trading off true positives against true negatives. The regimes that tend to produce the most similar results for black and non-black offenders, however, include simply using drug violations or reporting violations alone.

At least part of the relative performance of rules is attributable to the timing of violations. Fees and fines violations, for example, tend to accumulate later in the spell, when most individuals who are likely to reoffend have already done so (see Appendix Figure A18). As a result, the population at risk to fail to pay fees and fines is meaningfully positively selected. Timing is only partly responsible for the patterns in Figure V, however. It is straightforward to simulate the share of reoffenders who would break technical rules of each type at any point in their spell instead of the share who break rules before being rearrested. Producing a version of the figure with this quantity on the x-axis shows similar patterns (see Appendix Figure A19). In fact, for black men, fees and fines violations remain *negatively* correlated with reoffending risk: those who cannot pay are less likely to reoffend than those who can.

## VI CONCLUSION

This paper studies the primary way the US criminal justice system gives convicted offenders a second chance to avoid prison and get back to work: probation. After conviction, probationers return home but are subject to technical rules that forbid drugs and alcohol, require payment of fees and fines, and limit travel, among other constraints. Rule violators can be “revoked” and sent to prison, making probation an important driver of incarceration. Since black men are significantly more likely to break rules, probation also drives racial

disparities in prison exposure.

I use a 2011 reform in North Carolina that reduced prison punishments for technical rules to study whether rule violations are strong predictors of future reoffending and deter reoffending. I find that harsh punishments for rule breaking do little to encourage compliance, and that while rule violations are correlated with reoffending overall, they are significantly less predictive of future offending among black probationers. As a result, North Carolina’s reform closed racial gaps in revocation without affecting racial gaps in rearrests. Using a semi-parametric model of competing risks, I find that rules related to fees and fines are particularly poor tags of criminal risk and drive racial disparities.

Many states continue to use technical violations extensively today, as shown in Figure VI. This figure lists the top 20 US states ranked by the share of state prison admissions due to technical violations of probation and parole from data collected recently by the Council of State Governments Justice Center (CSG, 2019). Many states sit at well over 25%, including New York, Ohio, Mississippi, and South Carolina. Most of these states—those with blue bars—have no statutory limitations on which technical violations can lead to revocation. Those that do—the gray bars—have simple “hardship” exceptions for fees and fines violations. Reduced reliance on fees and fines for revocation is therefore likely to be an attractive reform for many jurisdictions

More broadly, my results show how ostensibly race-neutral policies—in this case the imposition of common sense rules designed to encourage desistance from crime and promote public safety—can generate large racial disparities not justified by the policies’ ultimate goals. The design and impact of rules and policies are a potentially powerful explanation for many observed racial disparities in criminal justice and beyond. Fortunately, correcting disparities due to disparate impact may be easier than changing biased decision makers’ behavior—be they police, judges, or prosecutors—since doing so is a matter of simply changing the rules themselves. The findings presented here provide clear evidence that such changes are politically feasible and can have large, persistent impacts on racial disparities.

## REFERENCES

- Abadie, Alberto.** 2002. “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models.” *Journal of the American Statistical Association* 97 (457): 284–292.
- Abbring, Jaap H., and Gerard J. Van Den Berg.** 2003. “The Nonparametric Identification of Treatment Effects in Duration Models.” *Econometrica* 71 (5): 1491–1517.
- Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan.** 2012. “Do Judges Vary in Their Treatment of Race.” *The Journal of Legal Studies* 41 (2): 1239–1283.
- Agan, Amanda, and Sonja Starr.** 2018. “Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment.” *The Quarterly Journal of Economics* 133 (1): 191–235.
- Aghion, Philippe, and Jean Tirole.** 1997. “Formal and Real Authority in Organizations.” *Journal of Political Economy* 105 (1): 1–29.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin.** 1996. “Identification of Causal Effects Using Instrumental Variables.” *Journal of the American Statistical Association* 91 (434): 444–455.
- Angrist, Joshua D., and Alan B. Krueger.** 1995. “Split-Sample Instrumental Variables Estimates of the Return to Schooling.” *Journal of Business & Economic Statistics* 13 (2): 225–235.
- Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson.** 2012. “The Impact of Jury Race in Criminal Trials.” *The Quarterly Journal of Economics* 127 (2): 1017–1055.
- Arnold, David, Will Dobbie, and Crystal S Yang.** 2018. “Racial Bias in Bail Decisions.” *The Quarterly Journal of Economics* 133 (4): 1885–1932.
- Arrow, Kenneth.** 1973. “Higher education as a filter.” *Journal of Public Economics* 2 (3): 193–216.
- Barnes, Geoffrey C., Jordan M. Hyatt, Lindsay C. Ahlman, and Daniel T.L. Kent.** 2012. “The effects of low-intensity supervision for lower-risk probationers: updated results from a randomized controlled trial.” *Journal of Crime and Justice* 35 (2): 200–220.
- Bayer, Patrick, and Kerwin Kofi Charles.** 2018. “Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940.” *The Quarterly Journal of Economics* 133 (3): 1459–1501.
- Becker, Gary S.** 1957. *The Economics of Discrimination*.. University of Chicago Press.
- Berk, Richard, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth.** 2018. “Fairness in Criminal Justice Risk Assessments: The State of the Art.” *Sociological Methods & Research* 0049124118782533.
- Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad.** 2019. “Incarceration, Recidivism, and Employment.” *Journal of Political Economy* Forthcoming.



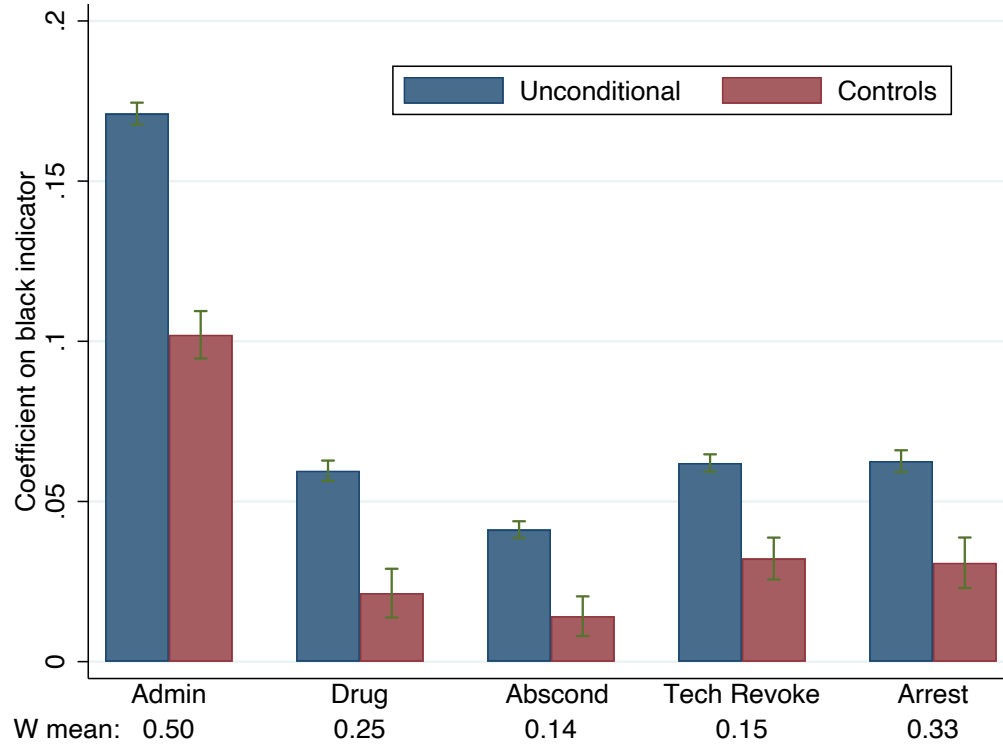
- Boyle, Douglas J., Laura M. Ragusa-Salerno, Jennifer L. Lanterman, and Andrea Fleisch Marcus.** 2013. "An Evaluation of Day Reporting Centers for Parolees." *Criminology & Public Policy* 12 (1): 119–143.
- Bushway, Shawn D., and Brian Forst.** 2013. "Studying Discretion in the Processes that Generate Criminal Justice Sanctions." *Justice Quarterly* 30 (2): 199–222.
- de Chaisemartin, Clément.** 2010. "A note on instrumented difference in differences." working paper.
- Chalfin, Aaron, and Justin McCrary.** 2017. "Criminal Deterrence: A Review of the Literature." *Journal of Economic Literature* 55 (1): 5–48.
- Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter.** 2018. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." NBER Working Paper No. 24441.
- Cohen, Mark A., Roland T. Rust, Sara Steen, and Simon T. Tidd.** 2011. "Willingness-To-Pay For Crime Control Programs." *Criminology* 42 (1): 89–110.
- Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi.** 2012. "Plausibly Exogenous." *The Review of Economics and Statistics* 94 (1): 260–272.
- Corbett-Davies, Sam, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq.** 2017. "Algorithmic Decision Making and the Cost of Fairness." In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17 797–806, New York, NY, USA: ACM.
- Cox, D. R.** 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society. Series B (Methodological)* 34 (2): 187–220.
- Cox, David R.** 1962. *Renewal Theory*. Methuen.
- Dobbie, Will, Jacob Goldin, and Crystal S. Yang.** 2018. "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *American Economic Review* 108 (2): 201–40.
- Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan.** 2018. "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India." *Econometrica* 86 (6): 2123–2160.
- Efron, Bradley.** 1988. "Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve." *Journal of the American Statistical Association* 83 (402): 414–425.
- Fryer, Roland G.** 2019. "An Empirical Analysis of Racial Differences in Police Use of Force." *Journal of Political Economy* 127 (3): 1210–1261.
- Hall, Michelle, Ginny Hevener, Susan Katzenelson, John Madler, Sara Perdue, and Rebecca Murdock.** 2015. "Justice Reinvestment Act Implementation Evaluation Report." Technical report, N.C. Sentencing and Policy Advisory Commission.
- Hall, Michelle, Ginny Hevener, Susan Katzenelson, John Madler, Sara Perdue, and Rebecca Wood.** 2014. "Justice Reinvestment Act Implementation Evaluation Report." Technical report, N.C. Sentencing and Policy Advisory Commission.

- Harding, David J., Jeffrey D. Morenoff, Anh P. Nguyen, and Shawn D. Bushway.** 2018. "Imprisonment and Labor Market Outcomes: Evidence from a Natural Experiment." *American Journal of Sociology* 124 (1): 49–110.
- Heckman, James J., and Bo E. Honoré.** 1989. "The Identifiability of the Competing Risks Model." *Biometrika* 76 (2): 325–330.
- Heckman, James, and Burton Singer.** 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52 (2): 271–320.
- Hennigan, Karen, Kathy Kolnick, Tian Siva Tian, Cheryl Maxson, and John Poplawski.** 2010. "Five Year Outcomes in a Randomized Trial of a Community-Based MultiAgency Intensive Supervision Juvenile Probation Program." technical report.
- Honoré, Bo E.** 1993. "Identification Results for Duration Models with Multiple Spells." *Review of Economic Studies* 60 (1): 241–46.
- Hunt, Heather, and Gene Nichol.** 2017. "Court Fines and Fees: Criminalizing Poverty in North Carolina." technical report, North Carolina Poverty Research Fund.
- Hyatt, Jordan M., and Geoffrey C. Barnes.** 2017. "An Experimental Evaluation of the Impact of Intensive Supervision on the Recidivism of High-Risk Probationers." *Crime & Delinquency* 63 (1): 3–38.
- Jannetta, Jesse, Justin Breaux, Hellen Ho, and Jeremy Porter.** 2014. "Examining Racial and Ethnic Disparities in Probation Revocation." washington, d.c. Urban Institute.
- Kaeble, Danielle, and Mariel Alper.** 2020. "Probation and Parole in the United States, 2017-2018." BJC Bulletin NCJ 252072, Bureau of Justice Statistics.
- Kaplan, E. L., and Paul Meier.** 1958. "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association* 53 (282): 457–481.
- Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan.** 2017a. "Human Decisions and Machine Predictions\*." *The Quarterly Journal of Economics* 133 (1): 237–293.
- Kleinberg, Jon, Snedhil Mullainathan, and Manish Raghavan.** 2017b. "Inherent Trade-Offs in the Fair Determination of Risk Scores." In *The 8th Innovations in Theoretical Computer Science Conference*, New York, NY, USA.
- Kline, Patrick, and Christopher R. Walters.** 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start\*." *The Quarterly Journal of Economics* 131 (4): 1795–1848.
- Kuziemko, Ilyana.** 2013. "How Should Inmates Be Released From Prison? An Assessment of Parole Versus Fixed Sentence Regimes." *Quarterly Journal of Economics* 128 (1): 371–424.
- Lyons, Christopher J., and Becky Pettit.** 2011. "Compounded Disadvantage: Race, Incarceration, and Wage Growth." *Social Problems* 58 (2): 257–280.

- Markham, James M.** 2018. “Monetary Obligations in North Carolina Criminal Cases.” technical report, University of North Carolina School of Government.
- Miller, T., M. Cohen, and B. Wiersema.** 1996. “Victim costs and consequences: A new look.” National Institute of Justice Research Report NCJ-155282, U.S. Department of Justice.
- Mueller-Smith, Michael.** 2015. “The Criminal and Labor Market Impacts of Incarceration.” *Working Paper*.
- Mueller-Smith, Michael, and Kevin T. Schnepel.** 2019. “Diversion in the Criminal Justice System.” working paper.
- Neal, Derek, and Armin Rick.** 2016. “The Prison Boom and Sentencing Policy.” *The Journal of Legal Studies* 45 (1): 1–41.
- Oaxaca, Ronald L., and Michael R. Ransom.** 1999. “Identification in Detailed Wage Decompositions.” *The Review of Economics and Statistics* 81 (1): 154–157.
- Pepin, Arthur W.** 2016. “The End of Debtors’ Prisons: Effective Court Policies for Successful Compliance with Legal Financial Obligations.” policy paper, Conference of State Court Administrators.
- Phelps, Edmund S.** 1972. “The Statistical Theory of Racism and Sexism.” *The American Economic Review* 62 (4): 659–661.
- Piehl, Anne Morrison, and Stefan F. LoBuglio.** 2005. “Does Supervision Matter.” In *Prisoner Reentry and Crime in America*, edited by Travis, Jeremy, and Christy Visher, New York: Cambridge University Press.
- Rambachan, Ashesh, and Jonathan Roth.** 2020. “An Honest Approach to Parallel Trends.” working paper.
- Reaves, Brian A.** 2013. “Felony Defendants in Large Urban Counties, 2009 - Statistical Tables.” State Court Processing Statistics NCJ 243777, Bureau of Justice Statistics.
- Rehavi, M. Marit, and Sonja B. Starr.** 2014. “Racial Disparity in Federal Criminal Sentences.” *Journal of Political Economy* 122 (6): 1320–1354.
- Robina Insitute.** 2016. “Probation Revocation And Its Causes: Profiles of State and Local Jurisdictions.” policy paper, University of Minnesota.
- Rose, Evan K., Jonathan Schellenberg, and Yotam Shem-Tov.** 2019. “The Effects of Teacher Quality on Criminal Behavior.” *Working Paper*.
- Rose, Evan K., and Yotam Shem-Tov.** 2019. “Does Incarceration Increase Crime?” *Working Paper*.
- Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek.** 2019. “PUMS USA: Version 9.0 [dataset].” Minneapolis, MN: IPUMS.
- Sakoda, Ryan.** 2019. “Efficient Sentencing? The Effect of Post-Release Supervision on Low-Level Offenders.” *Unpublished manuscript*.

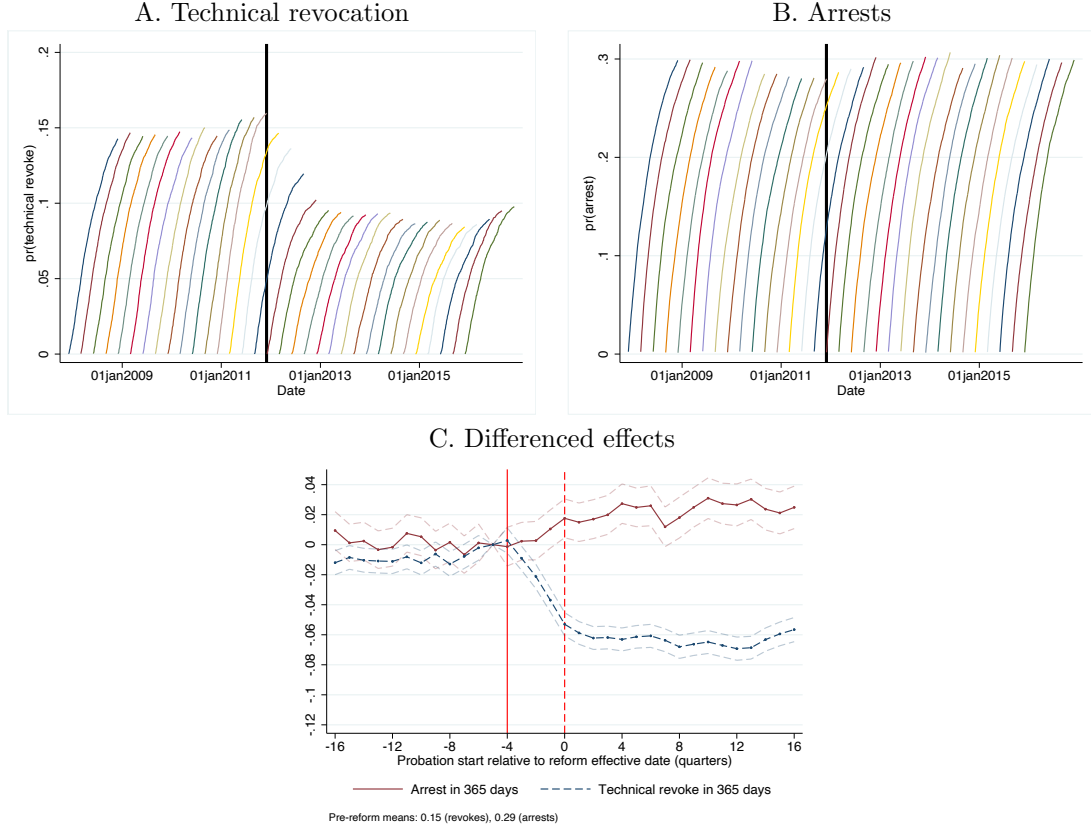
- The Council of State Governments Justice Center.** 2011. “Justice Reinvestment in North Carolina.” Technical report.
- The Council of State Governments Justice Center.** 2019. “Confined and Costly: How Supervision Violations are Filling Prisons and Burdening Budgets.” Technical report.
- Tsiatis, Anastasios.** 1975. “A Non-identifiability Aspect of the Problem of Competing Risks.” *Proceedings of the National Academy of Sciences* 72 (1): 20–22.
- Van Den Berg, Gerard J.** 2001. “Duration models: specification, identification and multiple durations.” In *Handbook of Econometrics*, edited by Heckman, J.J., and E.E. Leamer Volume 5. of Handbook of Econometrics, Chap. 55 3381–3460, Elsevier.
- West, Jeremy.** 2018. “Racial Bias in Police Investigations.” *Working Paper*.
- Western, Bruce, and Becky Pettit.** 2005. “Black-White Wage Inequality, Employment Rates, and Incarceration.” *American Journal of Sociology* 111 (2): 553–578.

FIGURE I  
RACIAL DISPARITIES IN PROBATION OUTCOMES



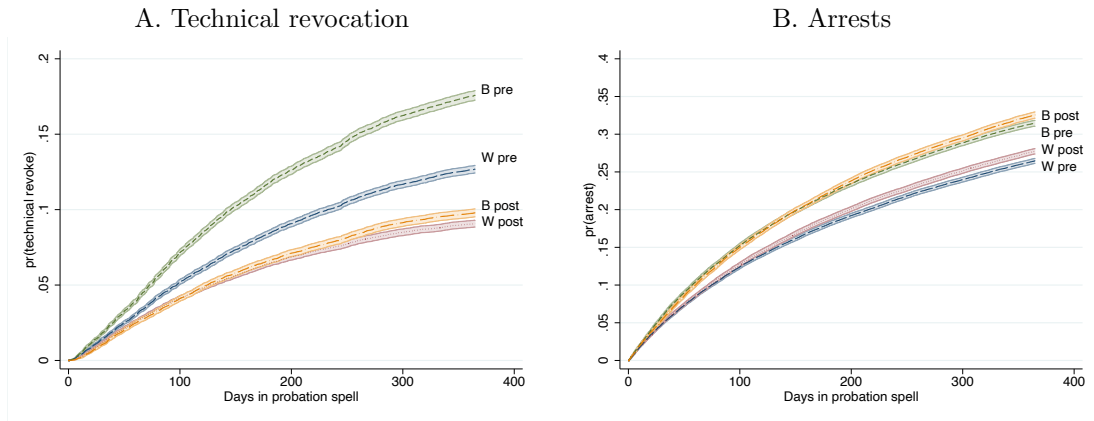
*Notes.* This figure presents regression results that include all supervised probationers starting spells in 2006-2010. W mean refers to the non-black mean of the dependent variable, which is an indicator for the relevant outcome occurring at any point in the spell. Admin includes violations such as non-payment of fees and fines. Drug includes drug-related violations. Absconding is fleeing supervision. Technical revocations are revocations without a preceding criminal arrest. Adjusted estimate is from an OLS regression with controls for gender, 20 quantiles of age effects, district fixed effects, fixed effects for the offense class of their focal conviction, a linear control for the length of the supervision spell, fixed effects for prior convictions and revokes, a linear control for previous incarceration duration, and the most recent math and reading standardized test scores (normalized to have mean 0 and standard deviation 1 in the full test-taker population) observed between grades 3 and 8. 95% confidence intervals are indicated by the whiskers atop each bar and are formed from standard errors clustered at the individual-level.

FIGURE II  
EFFECTS OF REFORM ON TECHNICAL REVOCATION AND CRIME



*Notes.* This figure plots effects of the 2011 JRA reform on technical revocation and arrests. Panels A and B include all supervised probationers starting their spells within four years of the reform. Each line represents a three-month cohort of probationers who start their spells where the line intersects the x-axis. The y-axis measures the share of this cohort experiencing the relevant outcome over the following year. That is, each line is the failure function for that cohort and outcome. Technical revocation is an indicator for having probation revoked with no intervening criminal arrest. Arrest is an indicator for a criminal arrest before revocation for any rule violations. Events are therefore mutually exclusive. Panel C plots mean one-year technical revocation and arrest rates for supervised probationers minus the same measure for unsupervised probationers. The same cohort definitions are used. Effects are normalized relative to the cohort starting four quarters before the reform, indicated by the solid red line. This is the last cohort to spend the full first year of their probation spells under the pre-reform regime. The dotted red line indicates the first cohort whose first year of probation falls completely post-reform. Dashed lines indicate 95% confidence intervals formed from standard errors clustered at the individual level.

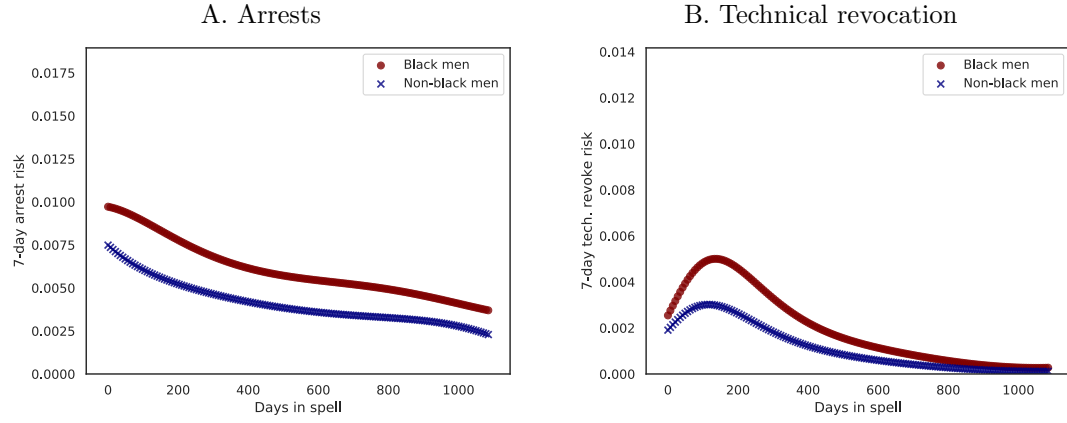
FIGURE III  
EFFECTS OF REFORM BY RACE



*Notes.* This figure plots effects of the 2011 JRA reform on technical revocation and arrests separately by race. It includes all supervised probationers starting their spells either 1-3 years before (pre) or 0-2 years after the reform (post). “B” refers to black probationers, while “W” refers to non-black. The y-axis measures the share of each group experiencing the relevant outcome over the first year of their probation spell. Technical revocation is an indicator for having probation revoked for rule violations with no intervening criminal arrest. Arrest is an indicator for a criminal arrest before revocation for any rule violations. Shaded areas reflect 95% confidence intervals formed using standard errors clustered at the individual level.

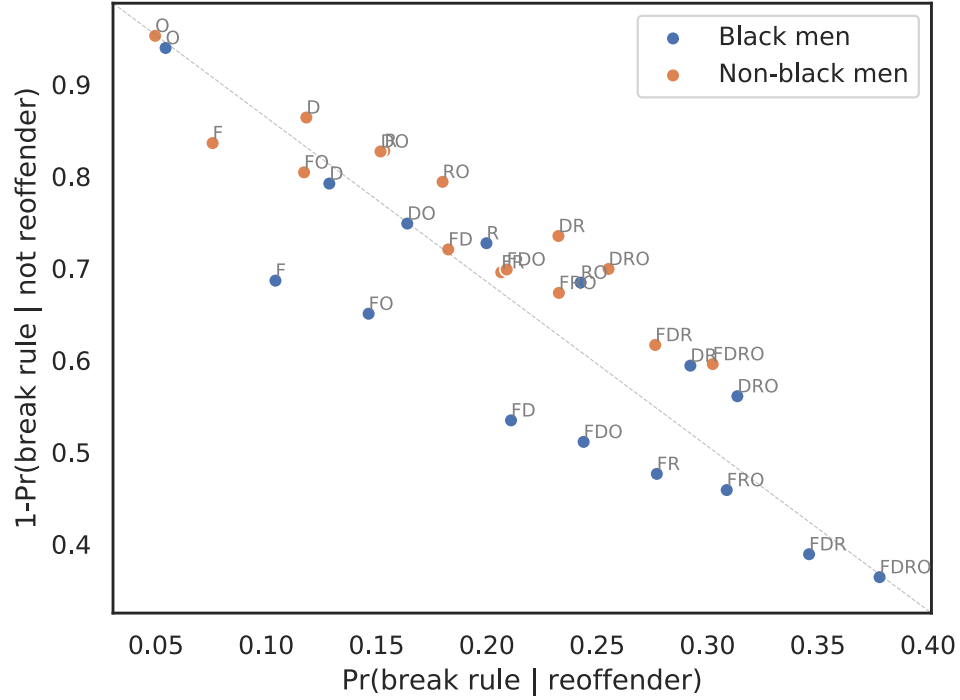


FIGURE IV  
AVERAGE HAZARDS FOR ARREST AND TECHNICAL REVOCATION



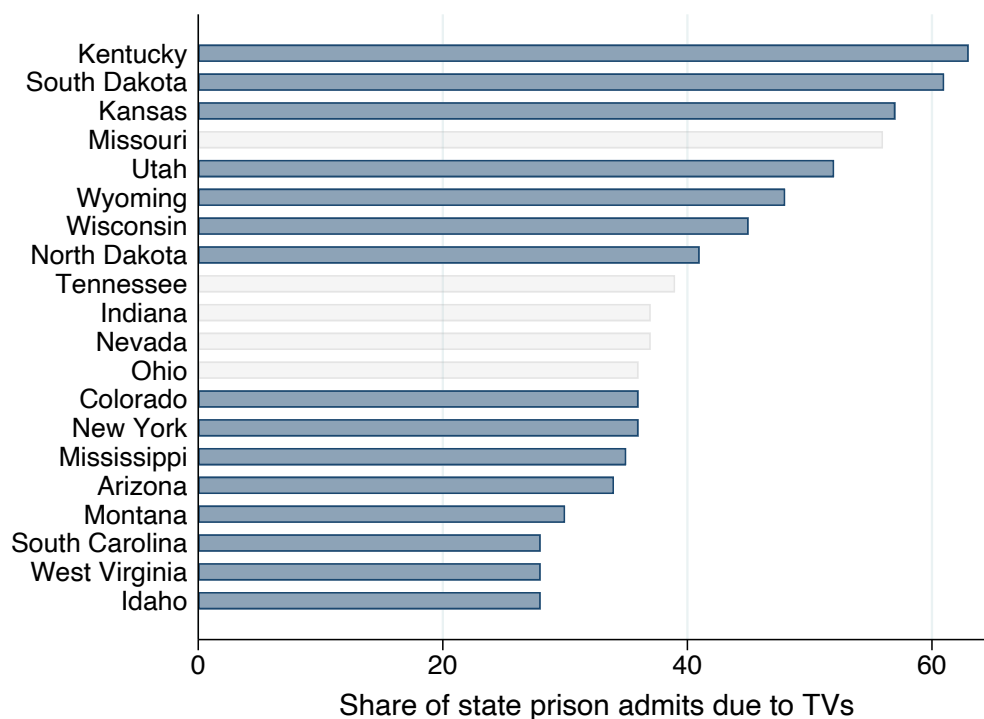
*Notes.* This figure plots average baseline weekly hazard rates for each outcome implied by estimates of the mixed logit competing risks model. The baseline hazard reflects the risk of each event for the *same individual* conditional on the event not happening previously. Hazards are calculated for an individual with mean levels of observables and averaged over the distribution of unobserved heterogeneity using estimates from finite mixture version of the model estimated with four types.

FIGURE V  
EFFICIENCY AND EQUITY OF TECHNICAL VIOLATION RULE TYPES



*Notes.* This figure plots estimates of the share of potential reoffenders over a three year period who break technical rules before they reoffend (x-axis) against the share of non-reoffenders who do not break technical rules. Estimates come from simulating the model estimated in Section VD using a different set of rules. Each point is labeled with a combination of “F” for fees / fines violations, “D” for drug / alcohol violations, “R” for reporting violations, and “O” for all other, reflecting the sets of rules enforced in the simulation. The points labeled “FDRO” therefore reflect the set of rules punishable with incarceration before the 2011 reform, and “R” reflect the set punishable afterwards. The dotted gray line starts at (1, 0) and has a slope of -1. This line reflects what would be achieved by randomly revoking a fraction of probationers at the start of their spells, which naturally would catch equal shares of reoffenders and non-reoffenders.

FIGURE VI  
TOP STATES BY SHARE OF PRISON ADMISSIONS DUE TO TECHNICAL VIOLATIONS



*Notes.* Figure plots the share of state prison admissions due to technical violations of probation and parole using data from the Council of State Governments Justice Center (CSG, 2019) for the 20 states with the highest shares. States with blue bars have no statutory limits on which technical violations can result in prison time, while states with gray bars restrict incarceration for failure to pay fees and fines when the defendant can demonstrate a financial “hardship.”

TABLE I  
DESCRIPTIVE STATISTICS

	Supervised (treated)			Unsupervised (control)		
	Mean	Sd.	p50	Mean	Sd.	p50
<b>Demographics:</b>						
Age at start	32.059	10.85	29.83	32.707	10.77	30.29
Male	0.738	0.44	1.00	0.732	0.44	1.00
Black	0.435	0.50	0.00	0.355	0.48	0.00
White	0.490	0.50	0.00	0.522	0.50	1.00
Other race	0.074	0.26	0.00	0.124	0.33	0.00
<b>Sentence:</b>						
Sup. length (m)	19.449	9.58	18.17	14.841	8.77	12.00
Felon	0.429	0.49	0.00	0.032	0.18	0.00
Misd.	0.318	0.47	0.00	0.502	0.50	1.00
DWI / DWLR	0.208	0.41	0.00	0.457	0.50	0.00
<b>Criminal history:</b>						
Crim. hist. score	2.059	2.97	1.00	0.988	1.76	0.00
Prior sentences	1.917	3.28	0.00	1.251	2.69	0.00
Prior inc. spells	0.860	2.22	0.00	0.497	1.74	0.00
<i>N</i>	708623			895090		
Individuals	531099			661103		

*Notes.* This table includes all supervised (treated) and unsupervised (control) probation spells beginning between 2006 and 2018. Felon, misdemeanor, and DWI / DWLR measure the most serious offense in the conviction that produced the probation sentence, with DWL / DWLR referring to driving while intoxicated and driving with license revoked. A small share of spells result from offenses with no classification. Criminal history score is a weighted sum of prior convictions used by North Carolina's sentencing guidelines. A prior misdemeanor conviction is typically worth 1 point, while a prior felony is worth two or more. Prior sentences refer to previous sentences to supervised probation or incarceration. Prior incarceration spells refers to previous incarceration in state prison.

TABLE II  
TOP 20 PROBATION VIOLATIONS

	Violation	Share of violations	Share of spells
	Any violation	1.000	0.618
1	Not paying fees	0.343	0.496
2	Not reporting	0.129	0.286
3	Positive drug test	0.085	0.184
4	Fleeing supervision	0.064	0.163
5	New misdemeanor charge	0.063	0.138
6	Treatment / program failure	0.061	0.156
7	Moving / job change without notifying	0.034	0.084
8	Not completing community service	0.033	0.102
9	Breaking curfew	0.028	0.065
10	No employment	0.023	0.059
11	New felony charge	0.019	0.040
12	Admitting drug use	0.009	0.023
13	No education / training	0.007	0.018
14	Travelling without permission	0.006	0.014
15	Possessing drugs	0.006	0.013
16	Electronic monitoring failure	0.004	0.010
17	Refuse drug test	0.003	0.008
18	Disobeying curfew	0.003	0.008
19	Possessing weapons	0.002	0.006
20	Contacting drug users	0.002	0.005
21	All others	0.162	0.558

*Notes.* This table includes all treated observations starting probation in 2006-2010. Share of violations measures share of all violation recorded over this period. Share of spells measures the share of probation spells with any violation of the listed type.

TABLE III  
DIFFERENCE-IN-DIFFERENCES ESTIMATES OF REFORM IMPACTS

A. All offenders				
	Technical revoke		Arrest	
	(1)	(2)	(3)	(4)
Post-reform	-0.00172*** (0.000273)	-0.00205*** (0.000288)	-0.00793*** (0.00167)	-0.00705*** (0.00159)
Treated	0.143*** (0.00103)	0.133*** (0.00102)	0.0316*** (0.00166)	-0.0155*** (0.00164)
Post-x-treat	-0.0532*** (0.00135)	-0.0530*** (0.00135)	0.0196*** (0.00242)	0.0194*** (0.00233)
<i>N</i>	546006	546006	546006	546006
Pre-reform treated mean	.149	.149	.287	.287
Accuracy			.369 (0.045)	.369 (0.063)
False negative rate			.936 (0.01)	.936 (0.01)
False positive rate			.056 (0.004)	.056 (0.004)
B. Non-black offenders				
Post-reform	-0.000522 (0.000317)	-0.000875** (0.000334)	-0.00693*** (0.00199)	-0.00666*** (0.00190)
Treated	0.122*** (0.00130)	0.112*** (0.00126)	0.0450*** (0.00209)	-0.000334 (0.00207)
Post-x-treat	-0.0356*** (0.00173)	-0.0360*** (0.00172)	0.0198*** (0.00304)	0.0179*** (0.00295)
<i>N</i>	328784	328784	328784	328784
Pre-reform treated mean	.127	.127	.265	.265
Accuracy			.556 (0.085)	.55 (0.081)
False negative rate			.93 (0.01)	.931 (0.01)
False positive rate			.025 (0.005)	.026 (0.005)
C. Black offenders				
Post-reform	-0.00387*** (0.000509)	-0.00412*** (0.000534)	-0.0118*** (0.00295)	-0.0112*** (0.00281)
Treated	0.167*** (0.00167)	0.160*** (0.00167)	-0.00496 (0.00274)	-0.0464*** (0.00268)
Post-x-treat	-0.0741*** (0.00215)	-0.0736*** (0.00214)	0.0228*** (0.00399)	0.0233*** (0.00383)
<i>N</i>	217222	217222	217222	217222
Pre-reform treated mean	.176	.176	.315	.315
Accuracy			.308 (0.053)	.309 (0.051)
False negative rate			.932 (0.01)	.932 (0.01)
False positive rate			.091 (0.007)	.091 (0.007)

*Notes.* This table includes all treated and untreated probation spells beginning 1-3 years before the reform and 0-2 years afterwards. Post is indicator for starting probation after Dec. 1, 2011, the date JRA reforms took effect. Technical revocation is an indicator for having probation revoked with no intervening criminal arrest. Arrest is an indicator for a criminal arrest before revocation for any rule violations. Demographic controls include five-year age bins and indicators for race and gender. Criminal history controls include fixed effects for criminal history points and prior sentences to supervised probation or incarceration. Controls are included in columns 2 and 4. Standard errors in parentheses are clustered at the individual level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

TABLE IV  
DECOMPOSITION OF RACIAL GAPS IN REVOCATIONS

	Overall rates		Decomposition	
	White	Black	Gap	Share of gap explained
Probability of technical revoke				
$Pr(R_i(0) = 1)$	0.039	0.082	0.043	100.0%
Distribution of risk				
$Pr(Y_i(0) = 1)$	0.313	0.376	0.063	9.8%
$Pr(Y_i(0) = 0)$	0.687	0.624	-0.063	-13.3%
True / false positive rates				
$Pr(R_i(0) = 1 Y_i(0) = 1)$	0.070	0.068	-0.002	-1.5%
$Pr(R_i(0) = 1 Y_i(0) = 0)$	0.025	0.091	0.066	104.9%

*Notes.* This table decomposes the difference in technical revocation between black and white probationers into the contributions of differences in reoffending risk and differences in the likelihood of revocation conditional on arrest risk. The decomposition applies to the population with  $R_i(1) = 0$  ( $\sim 90\%$  of the population). These are individuals who are not revoked for breaking rules even after the reform. Estimates are based on core difference-in-differences results without controls from Table III. The decomposition calculates the contribution of differences in risk using black targeting rates as baseline, and differences in targeting using white risk as baseline. The first row is -1 times the race-specific post-x-treat effect for technical violations. The second row is the sum of the constant, treat, and post-x-treat effects from difference-in-differences estimates for arrests. Both rows are re-scaled by 1 minus the sum of the constant, treat, and post-x-treat effects for technical violations, since this measures  $Pr(R_i(1) = 0)$ . The final two rows are calculated as described in the text. Appendix Section A4 provides complete details on how the decomposition is calculated.

TABLE V  
BEHAVIORAL RESPONSES TO REFORM

	Arrest		Any violation		Drug use		Fees and fines	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-reform	-0.00249 (0.0116)	0.000236 (0.0116)	-0.0205* (0.0101)	-0.0152 (0.0101)	0.0202 (0.0177)	0.0267 (0.0177)	0.00183 (0.0118)	0.00794 (0.0118)
<i>N</i>	152720	152720	152720	152720	152720	152720	152720	152720
Controls		Yes		Yes		Yes		Yes

*Notes.* Table reports estimates of Cox proportional hazard regressions using all supervised probation spells starting within one year of the reform. “Post reform” is a time-varying indicator for whether the period within the spell falls after Dec. 1, 2011. Each pair of columns considers the listed behavior as failure and the other behaviors as a source of independent censoring. In columns 1 and 2, for example, all arrests after an initial rule violation are ignored. If rule breaking and arrests are unaffected by the reform’s decrease in punishments for rule violations, then the populations at risk at each duration and measured hazards should also be unaffected. See Appendix Figure A4 for an illustration. Controls include five-year age bins, indicators for race and gender, and fixed effects for criminal history points and prior sentences to supervised probation or incarceration. All spells are censored at 365 days. Standard errors in parentheses are clustered at the individual level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .



TABLE VI  
TRIPLE DIFFERENCE ESTIMATES OF DIFFERENTIAL EFFECT ON BLACK OFFENDERS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Arrest	Revoke	Arrest	Revoke	Arrest	Revoke	Arrest	Revoke	Arrest	Revoke
Treat-x-post	0.0201*** (0.00304)	-0.0366*** (0.00175)	0.0129 (0.00784)	-0.0388*** (0.00496)	0.0192* (0.00786)	-0.0341*** (0.00496)				
Treat-x-post-x-black	0.00311 (0.00501)	-0.0394*** (0.00279)	0.00185 (0.00497)	-0.0375*** (0.00278)	-0.000708 (0.00504)	-0.0356*** (0.00284)	-0.00110 (0.00513)	-0.0352*** (0.00292)	-0.00283 (0.00563)	-0.0323*** (0.00311)
N	546006	546006	546006	546006	546006	546006	546006	546006	546006	546006
Demographics			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Criminal history					Yes	Yes	Yes	Yes	Yes	Yes
Probation district					Yes	Yes	Yes	Yes	Yes	Yes
Residence zipcode							Yes	Yes	Yes	Yes

*Notes.* This table includes all treated and untreated probation spells beginning 1-3 years before the reform and 0-2 years afterwards. Post is indicator for starting probation after Dec. 1, 2011, the date JRA reforms took effect. Demographic controls include five-year age bins and indicators for gender. Criminal history controls include fixed effects for criminal history points and prior sentences to supervised probation or incarceration. All controls are interacted with treatment, post, and treatment times post indicators. Standard errors are clustered at the individual level.

TABLE VII  
COST-BENEFIT ANALYSIS OF REFORM

	(1) $\Delta$ rev. \$	(2) $\Delta$ indir. \$	(3) Break-even	(4) Break-even (fel. only)	(5) Cost LB	(6) Cost UB
All	-633*** (25)	223 (119)	42,618*** (11,800)	104,818** (35,112)	16,757 (39,858)	177,546 (120,517)
Non-black	-423*** (33)	232 (129)	21,272* (10,246)	41,025* (20,473)	1,316 (41,011)	41,051 (124,545)
Black	-888*** (39)	309 (227)	47,976* (18,673)	179,263 (111,261)	30,206 (63,898)	331,070 (194,954)
Non-black men	-504*** (41)	226 (166)	27,115* (12,837)	45,461* (21,828)	-14,106 (44,689)	32,563 (140,121)
Black men	-1,004*** (48)	407 (301)	40,401* (18,363)	132,548 (84,530)	32,373 (69,830)	330,335 (211,819)

*Notes.* This table calculates the minimum mean social costs of arrests necessary for the state to “break-even” on changes in incarceration costs and arrest rates induced by the reform. Column 1 estimates the decrease in spending on revocation for technical violations per probationer due to the reform. Column 2 estimates the increase in spending on incarceration for new arrests. Columns 3 and 4 calculate implied break-even costs of an arrest for all arrests and for felony arrests only, respectively. Columns 5 and 6 report estimated increases in the costs of crime due to the reform when each arrest is assigned a dollar social cost using estimates from the literature. Includes all treated and untreated probation spells beginning 1-3 years before the reform and 0-2 years afterwards. Controls include five-year age bins, indicators for race and gender, and fixed effects for criminal history points and prior sentences to supervised probation or incarceration. Standard errors are clustered at the individual level.