Who Gets a Second Chance? Effectiveness and Equity in Supervision of Criminal Offenders

Evan K. Rose[*]

September 13, 2020

Abstract

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.

1   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 as to be holding a job. Recent research has studied racial disparities in decisions by police, judges, prosecutors, and juries (Fryer, 2019; Arnold, Dobbie and Yang, 2018; Rehavi and Starr, 2014; Anwar, Bayer and Hjalmarsson, 2012) and how arrest, conviction, and incarceration affect economic outcomes (Agan and Starr, 2018; Dobbie, Goldin and Yang, 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 way offenders are punished in the United States. Every year, more than 4.4 million convicted offenders are sent home on the condition that they obey strict technical rules. Breaking these rules, which 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.[1]

Technical rules, however, are the primary tools the corrections system uses to surveil convicted offenders and support their reintegration (Piehl and LoBuglio, 2005). Despite the costs, punishing rule breaking with incarceration—known as "revoking" supervision—may therefore be effective if violations are a strong indicator 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., 2017; Kleinberg, Mullainathan and Raghavan, 2017) and on differences in any behavioral responses to the threat of punishment.

This paper examines the effectiveness and equity of revocation for breaking technical rules 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 for rule breaking prior to the reform were instead permitted to remain in their communities or were subject to short periods of confinement. Measuring the resultant increases in crime 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 separately by race

---

[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).

allows me to assess equity by examining differences in targeting and deterrence effects between black and non-black probationers.

I begin with a reduced-form analysis of the 2011 reform. The analysis examines 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' outcomes are compared to those of individuals convicted of similar offenses and placed on unsupervised probation, an alternative punishment regime. Unsupervised probationers provide a useful control group because they are not subject to most technical rules and thus were unaffected by the reform. Their outcomes track the treated group's closely over the full pre-reform period. Results change little, however, in a simple pre-post analysis of the treated group alone instead.

Difference-in-differences estimates reveal that revocation for technical rule violations in the first year of a spell declined 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 offenders' revocation was nearly twice as large as its impact on non-black offenders'. As a result, racial gaps in revocation for technical rule violations were practically 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.

To interpret these results, I develop a simple empirical model that describes the relationship between two binary events: whether a probationer is revoked for technical violations and whether he is arrested for a crime. Casting the reform as in instrument in the classic Angrist, Imbens and Rubin (1996) framework, it is possible to estimate the probability offenders spared revocation would reoffend instead (Abadie, 2002). These reoffending rates can then be combined with other observed quantities to construct the accuracy, type-I (i.e., false positive), and type-II (i.e., false negative) error rates associated with using revocation for technical rule breaking as a tag for counterfactual reoffending. Critically, this approach requires that the reform affects reoffending only through the change in revocations. Validation tests 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 exclusion 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 who escaped 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 for rule breaking 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 rules catch a meaningful fraction of non-reoffenders and few potential reoffenders.

The effectiveness of revocation for technical violations as a tag for risk varies substantially by race. Roughly 50% of non-black probationers spared revocation were arrested, while among black probationers the arrest rate was only 30%. The implied accuracy of technical 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 rules.

Additional results suggest these race gaps reflect the disparate *impact* of ostensibly race-neutral rules rather than disparate *treatment* by those who enforce them. For example, there is no disparity in how black and non-black probationers are punished conditional on breaking the same technical rule. Moreover, technical rules for which officers have wide enforcement discretion and those for which violations are detected automatically both exhibit large race gaps. There is also no evidence of caseworker-probationer race match effects. This setting thus highlights the potential importance of how rules and policies are designed rather than how they are applied by practitioners for explaining racial disparities (Bushway and Forst, 2013; Neal and Rick, 2016). However, it is not possible to rule out that all probation officers and judges are uniformly biased against black offenders, which would not be detected in these tests. Substantial evidence of such disparate treatment due to either taste-based or statistical motivations (Becker, 1957; Phelps, 1972; Arrow, 1973) has been documented in many other criminal justice settings (Abrams, Bertrand and Mullainathan, 2012; Rehavi and Starr, 2014; Fryer, 2019; Arnold, Dobbie and Yang, 2018).

These findings are also consistent with previous work in Sakoda (2019). Using a similar difference-in-differences strategy, Sakoda (2019) finds that eliminating post-release supervision 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. Technical revocation for this population of offenders in Kansas therefore had similar racially disparate impacts to those observed in North Carolina for probationers.

Next, I extend the binary outcome model to examine racial differences in the effectiveness of technical violations at multiple horizons. The estimates show that black offenders are targeted more aggressively by revocation for technical violations regardless of whether and when they would be otherwise rearrested. There are especially large racial differences for probationers who would only be rearrested three years after starting probation or later (and possibly never). While black offenders are more likely to reoffend overall and to reoffend earlier in their spells, a decomposition shows that racial differences in reoffending risk explain less than 10% of racial differences in the likelihood of revocation for technical violations.

I use these results to conduct a partial cost-benefit analysis that compares the costs of incarcerating a technical rule breaker to the social costs of crime they would commit and any attendant punishments if allowed to remain free. The results show that for every $100 the state spends revoking technical rule breakers, it saves $30 it would have spent on incarceration for new offenses. To justify the state's use of revocation for technical rule breaking, the social costs of crime 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. Importantly, however, these calculations also assign no value to the impact of the reform on racial

3

disparities.

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 propensity to reoffend is tightly correlated with the propensity to break rules. Second, probationers may change their behavior in response to changes in rule enforcement, an effect ruled out in the reduced-form analysis. And finally, different types of rules may have very different effects. The reform however, impacted 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. In the model, probationers have latent risks of rearrest and revocation for breaking technical rules. Both risks evolve over the course of a spell, allowing for state dependence in behaviors. They also depend on observable characteristics such as age and criminal history and on unobserved probationer-specific 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 rather than ruling them out. To disaggregate rule types, an extension breaks the risk of rule-driven revocation into type-specific risks that also depend on both observable and unobservable factors. By estimating the model completely separately by race and gender, this approach can therefore capture rich differences in the relationship between rule breaking, reoffending, and rule enforcement across populations.

The estimates show that arrest hazards decline throughout the spell. Revocation risk, however, peaks roughly nine months into the spell. These patterns help explain the high error rates estimated in the reduced form: as a consequence of simple dynamic selection, most rule breakers have already revealed themselves to pose limited reoffending risk by virtue of not reoffending earlier in their spell. Nevertheless, individuals who are observably and unobservably more likely to reoffend are also more likely to break technical rules. The connection between rule breaking and reoffending risk is substantially weaker for black offenders, however. 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.

The estimates show limited evidence of behavioral responses to changes in enforcement. Weekly average latent arrest hazards are less than 0.1 p.p. higher after the change in policy. Violation behavior changes little as well, with very small estimated *decreases* in the risk of drug violations and failure to pay fees and fines. Probationers therefore do not appear responsive to weaker enforcement regimes. Moreover, estimated behavioral responses are similar across race groups, suggesting disparities in technical revocations are not justified by larger deterrent effects among black offenders. 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 future reoffenders who break technical rules and decrease the share of future non-reoffenders incarcerated 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 the 2011 reform directly addressed financial rules, it had large impacts on disparities within more limited impacts on crime. 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. Poorly 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 2. Section 3 lays out the empirical model. Section 4 presents the main results that analyze the 2011 reform. Section 5 estimates a competing risk model for probation violations and crime. Section 6 concludes.

2   Setting and data

2.1   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. The majority of these offenders—roughly 80%—are serving terms of probation, a period of community service ordered in lieu of a suspended incarceration sentence. 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 US residents living today 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.[2]

The size of the probation system reflects the popularity of probation as a criminal sentence. In the 75 largest counties in the US, 51% of felony defendants receive probation as part of their sen-

---

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

tences, 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.[3]

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 Statues §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 a roughly $30-50 monthly fee for supervision itself and repayment for any indigent defense provided, 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.[4] Of course, all probationers are also required to commit no new criminal offenses during their spell.

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 Insitute, 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 unsupervised probation, unsupervised probation tends to be reserved for misdemeanants and individuals convicted of driving while intoxicated or with a revoked license (descriptive statistics are presented in Table 1 and discussed further below). 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

---

[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]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.

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.[5] 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, which they often do. Over the 2000s, for example, probationers remanded to prison without a new criminal conviction accounted for ∼40% of state prison admissions.

Probation officers may have some discretion in whether to code a behavior such as drug 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. Results change little, however, when using alternative definitions of technical revocation, such as revocation for violations coded exclusively as non-criminal.

2.2   2011 reform

In 2011, North Carolina made major changes to the state's criminal justice system by passing the Justice Reinvestment Act (JRA).[6] 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).[7] Previously, judges could revoke for any technical violation, including non payment of fees and fines, not reporting, or failing drug and alcohol tests. In effect, the reform therefore barred revocation for drug and administrative violations but not for absconding, sharply reducing revocation rates.

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

---

[5]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.

[6]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 (CSG, 2011). Law makers passed the JRA in an effort to reduce these costs and lower projected correctional spending in the future.

[7]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.

no longer revoked for technical violations due to the reform received a CRV instead. I show below that for felony offenders, who serve their CRVs in state prisons where I can most reliably measure incarceration durations, JRA's changes amounted to meaningfully more lenient punishments for rule breaking. Offenders spared revocation due to the reform saw reductions of roughly 200 days of incarceration for technical violations accrued in the first year of their spells.[8]

Assessing the impact of CRVs on misdemeanor offenders is more difficult because they are typically served in local jails, where data on duration is not systematically collected.[9] Before June 2013, CRVs did not have to be served over a period of consecutive days (i.e., uninterrupted) (NC S.L. 2013-101). And in 2015, the North Carolina legislature eliminated CRVs all together for misdemeanor offenders (NC S.L. 2015-191), allowing me to examine how outcomes change after this alternative to revocation was eliminated. I return to how CRVs impact the interpretation of my results in Sections 3.3 and 4.4.

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 revocation authority and the introduction of CRVs 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 effects of the change to revocation policy while holding officer authority constant by looking in a relatively narrow window around December 1, which I do in robustness checks.

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.

2.3   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 the probation spell, the original convictions that resulted in the probation spell, and the probation officers assigned to the individual over the course of the spell. The data also record all violations (coded in dozens of unique categories), the probation officer's recommended response, and the ultimate disposition.

In addition to these records, I utilize data on all criminal court cases disposed from 2006 to the present provided by the North Carolina 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.[10] I use the AOC data to measure new criminal offenses, the length of any incarceration sentences meted out as a result, and criminal histories. I also use the AOC data to identify

---

[8]The average incarceration duration for felony probation revocations in 2010 was 221 days (median 188).

[9]The length of CRVs is also not recorded in the court datasets described below.

[10]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.

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 (NCERDC) and were linked to North Carolina criminal records for related work in Rose, Schellenberg and Shem-Tov (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 NCERDC 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.

2.4 Descriptive statistics

Descriptive statistics for the treated and control samples are provided in Table 1. 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. Despite these differences, I show below that control units' outcomes closely track those of treated units for many years leading up to the reform, supporting their use as a counterfactual.

As shown in Table 2, 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. 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.

Rather than work with the full list of detailed violation types, I categorize them into four groups: Drug related, administrative, absconding, and new crime. The top violations in each category are reported in Appendix Table A1. Drug related violations are predominately for failing a drug test, dropping out of a substance abuse program, or admitting to drug use. Administrative violations are predominately for non-payment of fees, not reporting, moving without permission, breaking curfew, failing to secure employment, etc. Absconding is a special violation issued when probation officers can no longer locate the offender. Arrest warrants are issued for absconders, and they are typically caught soon after. After the JRA reforms, offenders could only be incarcerated for new

crime or absconding violations. Beforehand, they could be revoked for any violation.

2.5 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.[11] Probation contributes to these patterns. Black offenders are more likely to face technical violations of virtually all types. These disparities are summarized in Figure 1, which reports the coefficients from regressions of a black indicator on an indicator for an event occurring within the probation spell using the North Carolina data. The blue bars report the coefficient when no additional controls are added, while the regressions underlying the red bars feature a battery of other controls, including covariates capturing demographics, geography, criminal history, and standardized math and reading test scores.[12] 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. Similar patterns have been documented in multiple other jurisdictions (Jannetta et al., 2014).

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 commit crime 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 in fact reflect that likely reoffenders are also likely to break technical rules.

3 Measuring effectiveness and equity

In this section, I provide a framework for assessing the effectiveness and equity of revocation when viewed as a simple tool for predicting socially costly behavior. In my context, technical rules—curfews, limitations on travel, and bans on drug and alcohol use, etc.—serve in part to identify offenders who are not committed to rehabilitation and thus likely to commit socially costly crimes. The same ideas, however, apply to other contexts, including bail setting (Kleinberg et al., 2017), parole release (Kuziemko, 2013), background screening, and rule breaking in non-criminal contexts, such as in classroom. 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.

---

[11]See Appendix Figure A1.

[12]Tables showing full regression results, including the effect of adding controls sequentially, are available starting with Appendix Table A2. Test scores are available due to related work in North Carolina described in Rose, Schellenberg and Shem-Tov (2019).

3.1 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 "spell," treating each person-spell observation as a separate unit to simplify exposition.[13]

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 reverse 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. Error rates provide alternative measures of effectiveness. When type-II error is close to zero, all individuals who would otherwise commit a crime are revoked for technical violations, making it easy to use rules to identify and imprison potential reoffenders.

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., 2017; Kleinberg et al., 2017; 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.[14] In what follows, I consider all three measures and pay particular attention to type-I and type-II errors because they are most closely connected to the concept of "disparate impact" discrimination in employment law.

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, Imbens and Rubin, 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. This final assumption requires that the entire change in reoffending after the reform is attributable to changes in revocation status. Imposing this exclusion restriction

---

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

[14]To see this, note that accuracy is related to error rates as:

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

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.

rules out offenders adjusting their reoffending behavior because probation overall has become a more lenient punishment. Such responses are potentially plausible. For example, offenders might use more drugs when failed drug tests are are punished less harshly, which could increase crime. I provide tests supporting this assumption in Section 4.3 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:

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

This quantity corresponds to the accuracy of revocation as a tag for counterfactual reoffending among the population with $R_i(1) < R_i(0)$. This population consists of offenders revoked for breaking rules targeted by the reform, namely drug and administrative violations. This technique therefore characterizes the accuracy of revocation for these types of violations specifically. Offenders who break rules unaffected by the reform, namely absconding violations, continue to be revoked afterwards and hence have $R_i(1) = R_i(0) = 1$.

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 also possible to estimate error rates. For example, type-II error rates are:

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

where $Pr(Y_i(0) = 1|R_i(1) = 0)$ and $Pr(R_i(1) = 0)$ can be easily estimated from their sample analogues. Error rates are identified for offenders with $R_i(1) = 0$, which empirically makes up $> 90\%$ of the full population for both grace groups. This population consists of those *not* revoked after the reform and therefore excludes absconders. As with accuracy, error rates thus characterize the effectiveness of drug and administrative rules specifically. In Section 5, 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 readily 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:

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

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

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

Thus the total difference is comprised of a component driven by differences in risk ($Pr(Y_i(0) = 1)$

12

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.

3.2 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. A simple extension to the one-period model allows me to examine racial differences in the effectiveness of probation across all horizons.

Specifically, let $Y_i \in \{0, 1, \ldots, \infty\}$ denote how many days after starting probation an individual reoffends, with ∞ indicating never. Let $R_i \in \{0, 1, \ldots, \infty\}$ measure how many days into a spell a probationer is technically revoked. As before, index potential revocation by the reform as $R_i(0), R_i(1)$, and index potential reoffending by R_i so that $Y_i = Y_i(R_i)$. It is then possible to estimate k-specific accuracy and error rates that measure the impacts of technical revocations at each horizon k:

$$\text{Accuracy} = Pr(Y_i(R_i(1)) = k | R_i(0) < k, R_i(1) > k)$$
$$\text{Type-I error} = Pr(R_i(0) < k | Y_i(R_i(1)) > k, R_i(1) > k)$$
$$\text{Type-II error} = Pr(R_i(0) > k | Y_i(R_i(1)) = k, R_i(1) > k)$$

Appendix Section A1 describes the additional assumptions required to estimate these objects and provides a derivation. Here, accuracy measures the likelihood that offenders revoked for technical rule violations prior to k would have otherwise reoffended at time k. Type-I error measures the likelihood that non-reoffenders by time k are revoked for technical violations. Type-II error measures the likelihood that reoffenders at time k are not revoked prior to k. The conditioning on $R_i(1) > k$ captures the fact that the reform did not completely eliminate revocation, so we can only estimate k-specific accuracy and error rates for the population not revoked by time k under the post-reform regime. To summarize the overall impact of differences in targeting vs. differences in reoffending risk, the decomposition in Equation 2 can also be extended by summing over all k instead simply the binary indicator in the one-period model.

3.3 Connection to empirical setting

As noted previously, some offenders no longer revoked due to the reform may have been confined for at most 90 days under a CRV, the new punishment introduced by the JRA legislation, instead. 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 3.1 estimates clearly

interpretable accuracy and error rates.[15] $Y_i(0)$, 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, I also consider several exercises that examine any direct effects of CRVs, including estimating effects after they were eliminated in 2015 and measuring the reform's impact on incarceration due to either CRVs or revocation by race.

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.

4   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 present estimates from the dynamic model over a three-year time horizon and several other extensions and robustness checks.

4.1   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. For each probationer, I measure which event occurs first (if any) and the time to the event. I then calculate the share of probationers technically incarcerated and the share arrested over the course of their spell.

Figure 2 plots the raw data for these two outcomes in Panels A and B, respectively, for three-month cohorts of supervised probationers. Each line represents a 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 (i.e., the cohort's failure function). All cohorts start their spells within four years of the reform's effective date, which is marked with the black solid line. 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 imprisoned for technical violations. Similar shares experience the same fate in each cohort for the next 12 quarters.

Cohorts beginning probation within a year of the reform, however, begin to see reductions in

---

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

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 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. Revocation then stabilizes for the next several years.

The large decrease in revocation means many more probationers had the opportunity to commit crimes instead of being revoked. 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. A meaningful share—roughly 30%—of probationers spared revocation in the first year of their probation spells were arrested instead.

This simple interrupted time series analysis may be misleading if selection into probation changed as a result of the reform or if changes in aggregate crime coincided with its implementation. 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 and I cannot reject the null the predicted 1-year crime rates are identical for spells starting in the year before vs. after the reform. Appendix Table A10 shows that predicted reoffending rates and other core covariates are also not trending differentially in treated vs. control units. 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. 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.

4.2   Difference-in-differences estimates

To account for potential time-varying confounders, I use a difference-in-differences approach that compares supervised probationers' outcomes to unsupervised probationers'. Panel C of Figure 2 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).[16] Specifically, it plots estimates of β_l^T from the linear regression:

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

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 β_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.

---

[16]The raw rates for unsupervised probationers are presented in Appendix Figure A7.

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 reform and 0-2 years afterwards, thus using two years of pre/post data while omitting cohorts whose first year of probation was interrupted by the reform and were therefore only partially affected.[17] These results are presented in Panel A of Table 4. The estimated effect on revocation is 5.5 p.p and easily distinguishable from zero at conventional confidence levels. The increase in arrests is roughly 2 p.p. Thus, over this one-year horizon 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. Moreover, the small coefficients on the post indicators show that over this narrow window, results would be similar if only treated units were 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 instead commit crimes, 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 5, where I estimate arrest and revocation hazards directly and show that they are highly correlated across individuals.

In the last two rows of Panel A, I use these results to estimate false positive and false negative rates, treating the full first year of the spell as a single period. Specifically, $Y_i = 1$ if an individual commits a crime in the first year of probation and is zero otherwise. The estimated false negative rate shows that just 6.5% 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 5.8% 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 that follow. 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 3 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

---

[17]I use these partially affected cohorts in estimation of the dynamic model that follows.

the race gap is reduced to less than 1%.

Because many more black offenders were spared revocation, one might expect crime in the black population to increase more than in the non-black population after the reform. Panel B of Figure 3 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 4 imply that the arrest rate among non-black offenders who would have been revoked for technical violations but for the reform is above 55%. However, the corresponding figure among black offenders is only 30%. 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 are revoked.

These results are remarkably similar to those in Sakoda (2019)'s evaluation of a reform to post-release supervision in Kansas.[18] In 2000, Kansas eliminated post-release supervision 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., but had no impact on new felony convictions that resulted in prison time. Thus incarceration for technical violations likely has similarly racially disparate impacts on probation as in post-release supervision or parole.[19]

Table 5 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 for technical violations.[20] 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. This implies that the bulk of differences in revocation are in fact driven by differences in how non-offenders are targeted. The last row of the table confirms this, showing that differences in false positive rates explain 105% of the aggregate gap.

The appendix contains several robustness exercises and extensions of these main results. Appendix Table A11 drops the control group and estimates single-difference effects. The patterns change little, with accuracy for black offenders roughly half that of non-black offenders and type-I error rates roughly 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.[21] Finally, Appendix Table A15 shows that the increase in crimes by crime 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, and a larger share of the total increase is accounted for by traffic related

---

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

[19]In North Carolina, parolees are supervised by the same officers as probationers and are required to follow a similar set of rules.

[20]Appendix Section A4 provides complete details on how the decomposition is calculated.

[21]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.

offenses. 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.

4.3 Testing exclusion

Using the reform as an instrument for revocation requires that offenders do not respond to the change in incentives by increasing their reoffending. Such behavioral responses find little support in the data. The risk of breaking a rule (regardless of the ultimate punishment) does not change after the reform takes effect. Nor do arrest hazards. Table 3 demonstrates this by estimating a post-reform effect in Cox proportional hazards models for arrests and rule-breaking. When studying arrests, 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 (see Appendix Figure A4 for a graphical illustration).

Analogous regressions can be estimated to test whether rule violations themselves increase after the reform. The results show no change in any behaviors. While perhaps surprising, these results are consistent with a series of randomized 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, Hansen and Rossi (2012) and Rambachan and Roth (n.d.). Racial differences in accuracy persist so long as violations of exclusion are not large and opposite signed for each race group.

4.4 Impacts of CRVs

As noted earlier, JRA introduced the option to impose short periods of confinement (CRVs) as a substitute for revocation for technical violations. Reoffending rates captured in the preceding analysis therefore reflect crime under this alternative policy rather than no confinement at all; racial differences in accuracy and error rates show that revocation had disparate impacts relative to outcomes under this alternative, as discussed in Appendix A2. It is possible, however, that reoffending outcomes would be different if there were no option to impose CRVs.

Several exercises suggest that CRVs do not explain the estimated racial disparities, however. 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 A12 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 on average respectively.

The larger decreases in incarceration for black offenders make their smaller observed increases in offending even more surprising.[22] Second, CRVs were eliminated for misdemeanor offenders in 2015. Appendix Table A13 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 A14 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.

4.5 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% 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 5.4 uses a competing hazard model to do so.

4.6 Triple-difference estimates

The previous results demonstrate that revocation for technical violations 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:

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

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. β_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

---

[22]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.

black and non-black offenders would be identical. Including X_{is} allows me to make this comparison after adjusting for observable characteristics. For example, the reform may have also had different impacts on men and women. When including a gender indicator in X_{is}, specification 4 tests whether racial differences in the impact of the reform still persist after accounting for differences in gender shares between the two race groups.

Table 6 reports estimates of β_7, labeled "treat-x-post-x-black", and β_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'.[23] 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. Even after adjusting for all these factors, black offenders continue to see substantially larger decreases in incarceration but no different changes in reoffending rates.

These results need not imply that *race itself* drives the differential impact of probation's technical rules. As argued in Section 4.9 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 6 shows that such differences are not easily explained with observable characteristics, including reasonable proxies for income such as residential 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.

4.7 Cost-benefit analysis

When the state revokes an offender for technical violations, it pays on average \$100 a day to do so.[24] If the state instead opts to leave the offender in the community, she may then commit a crime and be sentenced to incarceration as a result. The social value of technically revoking individual i can thus be written as:

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

where J_i is the cost of the technical jail/prison spells, R_i and Y_i, as before, are indicators for rule-driven revocation and offending, $U(Y_i)$ represents the social cost of this crime, and J_i' represents

---

[23]The post-x-treat coefficients reported here are identical to the post-x-treat estimates in Panel B of Table 4 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.

[24]2018 average daily cost per inmate for the North Carolina Department of Public Safety (https://www.ncdps.gov/adult-corrections/cost-of-corrections). Marginal costs of incarceration may be lower, but the JRA reform lead to substantial declines in incarceration that enabled the Department of Public Safety to close several facilities (Hall et al., 2015). Supervision costs roughly \$5 a day in 2018.

the total cost of incarceration as a result of the new crime, 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:

$$E[U(Y_i)|R_i = 0, Y_i = 1] = \frac{\triangle E[-J_i \cdot R_i] - \triangle E[(1 - R_i)J_i']}{\triangle E[Y_i]} \tag{6}$$

This exercise asks what the *minimum* social cost of crime would be to justify the state's use of revocation for the drug and administrative rules impacted by the reform. The numerator captures the change in net incarceration costs—spending on rule-driven incarceration minus spending on crime-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.[25]

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).[26] 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 potential criminals' offending calculus. Nor is there any change in technical violation behavior after the reform, including for payment of fees or fines.[27] 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 crime. 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 7. 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 \$680 per probationer on average. The second column reports the increase in costs of incarceration

---

[25]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.

[26]See Appendix Table A25 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.

[27]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).

attributable to new crimes committed in the first year of a spell. This is relatively close to zero because the majority of new crimes after 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 30 cents it would have spent on prison costs anyway.

Column 4 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. 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 7 repeats the same exercise for various sub-populations. The second and third rows, which compare black and non-black probationers, provide a concise summary of the degree to which drug and administrative violations target black offenders more aggressively. The decrease in spending on rule-driven incarceration in the black population is roughly twice as large, while increases in the costs of incarceration attributable to new crimes are only slightly larger. Combined with similar 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 7 shows that similar but more extreme patterns hold when considering black and non-black men.

4.8   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 3.2. 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 A8 plots type-II error rates for black and non-black offenders over a three-year horizon.[28] 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

---

[28]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.

aggregate disparity. Black revocation for drug and administrative violations would have been virtually identical to non-black revocation if they had similar conditional likelihoods to non-black offenders but their reoffending distribution were left the same.

4.9 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 these cases, however, 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 the 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 A9, controlling for assigned officer has no measurable impact on the black effect for technical violations. Relatedly, as Appendix Table A9 also shows, 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.[29] 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 Department of Public Safety'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 A8. 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.

---

[29]For drug violations, black officers treat black offenders slightly *more* harshly on average. There is no same-race effect in revocations overall, however, and small negative same race effect for technical revocations.

5 Competing hazards analysis

The previous results demonstrate that the technical rules affected by North Carolina's 2011 reform proxy for latent reoffending risk, but target black offenders substantially more aggressively. This analysis, however, leaves several important questions unanswered. For example, the timing of arrests and rule violations are potentially important drivers of effectiveness and may contribute to disparate impacts. Probationers may also respond to weaker punishment regimes by increasing rule violations or criminal activity, an effect ruled out by the exclusion restriction imposed previously. Different rules may also have different impacts both overall and on specific race groups.

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 technical violations, revocations, and reoffending that provides a solution by explicitly accounting for both observed and unobserved heterogeneity in behaviors.

5.1 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 hazards for individual i in period t of their sth probation spell are given by:

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

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

$\theta_0^Y(t)$ and $\theta_0^R(t)$ are unrestricted baseline hazards for each outcome shared by all individuals.[30] 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 a 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).[31] The two outcomes' hazards can be correlated through observables through β^Y and β^R and unobserved heterogeneity through U_i^R and U_i^Y. If offenders with high U_i^Y have high U_i^R as well, then even among observably equivalent offenders those more likely to be arrested are also more likely to break technical rules, and vice versa. With knowledge of $\theta_0^Y, \theta_0^R, \beta^Y, \beta^R$ and the

---

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

[31]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.

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 an indicator for whether period t falls before or after the 2011 reform in X_{ist}, one can also measure how each hazard responded to the change in policy. The coefficient on a post-reform indicator in the hazard for R_{is}^* should be large and negative, because the reform made revocation for rule breaking much less likely. The coefficient on a post-reform indicator in the hazard for Y_{is}^*, however, measures behavioral responses in reoffending to the reform and could take any sign.

Identification of $\theta_0^Y, \theta_0^R, \beta^Y$, and β^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.[32] 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 crime 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. As I show below, however, there is little evidence for increases in latent reoffending risk after the reform, consistent with my assumptions in the reduced form analysis.[33]

5.2   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.[34] 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 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.[35]

---

[32]Individuals have multiple spells because they frequently reoffend and are re-sentenced to probation. As shown in Table 1, there are 1.33 spells per person in the treated sample.

[33]Appendix Section A6 provides additional details on identification.

[34]In this sense, although the unobserved heterogeneity terms are treated as random effects, they are "correlated" random effects for the observables of interest (i.e., race).

[35]Adding additional types increases the likelihood but does not change any of the core conclusions discussed below. As

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.[36] 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:

$$\begin{pmatrix} U_i^Y \\ U_i^R \end{pmatrix} \sim N(\alpha, \Sigma) \tag{9}$$

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 Σ. Estimation of both versions is conducted in Python using the Boyd-Fletcher-Goldfarb-Shanno algorithm and the analytic gradient, which is straightforward to compute. Expectation Maximization algorithm estimation of the mixture version yields identical results, but is significantly slower.

5.3   Results

The estimates of this model support the conclusions from the previous analysis and reveal several important new insights. Figure 4, 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 shows that unobserved heterogeneity is an important feature of the data. Among black men, for example, the lowest reoffending risk type comprises 12% of the population and has a 3.5 log point lower weekly odds of offending then the highest risk type, which makes up 8% of the population.[37] Unobserved arrest risk has a strong correlation with unobserved revocation risk. Black men with the highest reoffending risk, for example, have 1.04 log point higher weekly odds of technical revocation than those with the lowest reoffending risk. Low-risk non-black men have even lower risk of technical revocation, with 6% 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

---

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.

[36]Additional details are provided in Appendix Section A6.

[37]Full parameter estimates are reserved for Appendix Tables A19 and A20.

offenders have already "selected out" of the pool still on probation. In this sense, time elapsed is one of the most potent signals of a probationer's future reoffending risk.

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 impact of all technical violations and not just those impacted by the 2011 reform. The pattern remains the same, however. Appendix Figure A9 demonstrates this by plotting $Pr(R_i^* < Y_i^* | Y_i^* = k)$ 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.[38]

Estimates of the effect of the reform on hazards show large impacts on the odds of revocation, which is 0.51 log points and 0.4 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 A13 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.

5.4   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 about how breaking different rules predicts reoffending risk. 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 may be more strongly connected to reoffending risk than the later. This variation is difficult to use in the reduced form because some offenders break rules and go unpunished only to be revoked for other violations later in their spell.[39] I cannot observe whether these individuals would have otherwise gone on to reoffend, making it difficult to measure the accuracy and error rates of specific types of rules. The hazard formulation used here accounts for this censoring.

---

[38]Part of this racial difference in targeting is driven by differences in observed characteristics, such as age and criminal history, while the remainder is driven by unobserved heterogeneity. Appendix Figure A14 shows that unobserved heterogeneity is responsible for most of the bias. This plot holds each race group's covariates fixed at the sample mean. The patterns change little.

[39]See Section 4.5.

Specifically, I decompose the latent risk of technical revocation into two components:

$$Pr(R_{is}^* = t | R_{is}^* \geq t) = Pr(V_{ist}^k = 1 | R_{is}^* \geq t) Pr(I_{ist} = 1 | V_{ist}^k = 1, R_{is}^* \geq t) \qquad (10)$$

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, allowing me to capture the variation discussed above. I model both components using a similar logit structure:

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

$$Pr(I_{ist} = 1 | V_{ist}^k = 1, X_{ist}, U_i^{V^k}, R_{is}^* \geq t) = \Lambda \left(\theta_0^I(t) + X_{ist}'\beta^I \right) \qquad (12)$$

The θ_0 terms terms describe how the risk of type k rule violations and revocation evolves within a spell. The relationship between β^Y and β^{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.[40] 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 12.

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. 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 A23 and A24. These estimate show substantial evidence of unobserved heterogeneity and state dependence as well. The estimates also continue to show economically small increases in the risk of reoffending as a result of the change in policy. The risks of violations also change little. Drug violations and 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

---

[40]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.

offenders' reoffending or rule-breaking behavior.

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 5 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 before doing so. In other words, the x-axis measure share of would-be reoffenders caught by technical rules, or the true positive rate. The y-axis plots the share of non-reoffenders over the same period who do not violate any rules. 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.[41]

The revocation 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. Consistent with the previous analysis, the regime using all rules ("FDRO") that corresponds to the pre-reform policy is roughly as likely to catch black reoffenders as non-reoffenders. This pre-reform regime does substantially better than this random guess frontier for non-black offenders.

Figure 5 illustrates several other interesting features of technical rules. First, using rules related to fees and fines is almost always dominated by not doing so for both race groups. For black offenders, for example, regimes that use fees and fines lie below and to the left of regimes that do not. Many sets of rule dominate using fees and fines alone. Switching from enforcing fees to enforcing drug violations, for example, would result in catching 2-3 p.p. more would-be reoffenders and imprisoning 12 p.p. fewer non-reoffenders Adding fees and fines to many regimes for black offenders in fact generates *worse* outcomes than a random guess, pulling outcomes within the frontier denoted by the gray line. Eliminating revocation for fees and fines violations thus offers a clear improvement over the current status quo.[42] North Carolina's reform achieved some of this impact by addressing this violation category.

Second, 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 innocents. Some rules, however, appear to be particularly unfair to black offenders. While fees and fines, for example, reduce the effectiveness of almost all regimes for non-black offenders, the decreases in true negative and true positive rates when using them in combination with other rules are smaller than for black offenders. Hence, dropping revocation for fees and fines rules thus 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 guessing. For non-black offenders, the pre- vs. post-reform shift appears to largely fall along possibility frontier.

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 increases in the share of would-be reoffenders incarcerated for increases in the share of non-reoffenders revoked. The

---

[41]Other rules include violations rarely charged, such as failing to pursue vocational training are contacting a victim.

[42]Ignoring impacts on collection, as discussed above.

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. The optimal technical rule regime depends on how policy makers assign benefits to catching would-be offenders and costs to revoking non-reoffenders. If the former is assigned more weight than the later, combinations of drug, reporting, and all other rules will be preferred. If the latter is assigned more weight, on the other hand, relying on smaller subsets of rules will be optimal.

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 A15). 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 5, 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 A16). 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.

6   Conclusion

This paper studies the largest part of the US community supervision system and the primary way the 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 crime and deter reoffending and to examine how their predictive power and deterrence effects differ across racial groups. I find 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 for breaking technical rules without affecting racial gaps in crime. Using a semi-parametric model of competing risks, I find that rules related to fees and fines are particularly poor predictors of future crime and drive racial disparities. I also find harsh punishments for rule violations have negligible deterrence effects that do not differ by race.

Many states continue to use technical violations extensively today, as shown in Figure 6. 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). In Kentucky, South Dakota, Kansas, Missouri, Utah, and Wyoming, technical violations among probationers and parolees account for more than 40% of prison admissions. Many other 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 tech-

nical violations can lead to prison time. 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. Indeed, related reforms have become increasingly popular in other areas of the criminal justice system, such as California's recent efforts to eliminate cash bail for pre-trial detention.

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 both 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., and Alan B. Krueger. 1995. "Split-Sample Instrumental Variables Estimates of the Return to Schooling." *Journal of Business & Economic Statistics*, 13(2): 225–235.

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.

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.

Chalfin, Aaron, and Justin McCrary. 2017. "Criminal Deterence: 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." *KDD '17*, 797–806. New York, NY, USA:ACM.

Cox, David R. 1962. *Renewal Theory.* Methuen.

Cox, D. R. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2): 187–220.

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." 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." 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, and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, 52(2): 271–320.

Heckman, James J., and Bo E. Honoré. 1989. "The Identifiability of the Competing Risks Model." *Biometrika*, 76(2): 325–330.

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.

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 Jeremey Porter. 2014. "Examing Racial and Ethnic Disparities in Probation Revocation." Urban Institute Washington, D.C.

Kaeble, Danielle, and Mariel Alper. 2020. "Probation and Parole in the United States, 2017-2018." Bureau of Justice Statistics BJC Bulletin NCJ 252072.

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. 2017. "Human Decisions and Machine Predictions*." *The Quarterly Journal of Economics*, 133(1): 237–293.

Kleinberg, Jon, Snedhil Mullainathan, and Manish Raghavan. 2017. "Inherent Trade-Offs in the Fair Determination of Risk Scores." 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.

Miller, T., M. Cohen, and B. Wiersema. 1996. "Victim costs and consequences: A new look." U.S. Department of Justice National Institute of Justice Research Report NCJ-155282.

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." Conference of State Court Administrators Policy Paper.

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.* , ed. Jeremy Travis and Christy Visher. New York:Cambridge University Press.

Rambachan, Ashesh, and Jonathan Roth. n.d.. "An Honest Approach to Parallel Trends." Working Paper.

Reaves, Brian A. 2013. "Felony Defendants in Large Urban Counties, 2009 - Statistical Tables." Bureau of Justice Statistics State Court Processing Statistics NCJ 243777.

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." University of Minnesota Policy Paper.

Rose, Evan K., and Yotam Shem-Tov. 2019. "Does Incarceration Increase Crime?" *Working Paper.*

Rose, Evan K., Jonathan Schellenberg, and Yotam Shem-Tov. 2019. "The Effects of Teacher Quality on Criminal Behavior." *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."

The Council of State Governments Justice Center. 2019. "Confined and Costly: How Supervision Violations are Filling Prisons and Burdening Budgets."

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.* Vol. 5 of *Handbook of Econometrics*, , ed. J.J. Heckman and E.E. Leamer, Chapter 55, 3381–3460. Elsevier.

West, Jeremy. 2018. "Racial Bias in Police Investigations." *Working Paper.*

Figures

Figure 1: Racial Disparities in Probation Outcomes
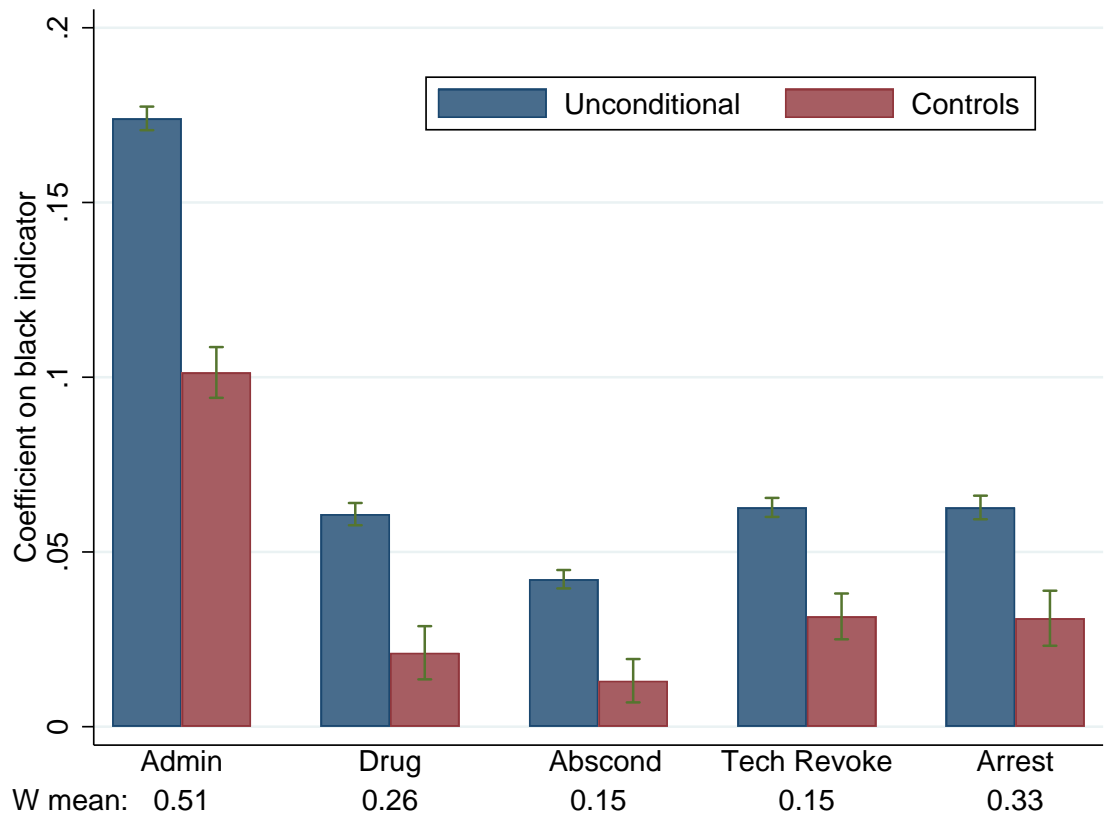
Notes: Regressions include all supervised probationers starting spells in 2006-2010. W mean refers to the non-black mean of the dependent variable, which is an indicator for the relevant outcome occurring at any point in the spell. Admin includes violations such as non-payment of fees and fines. Drug 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 2: Effects of Reform on Technical Revocation and Crime

A. Technical revocation

B. Arrests

C. Differenced effects

Arrest exit in 365 days — — — Technical revoke in 365 days

Pre-reform means: 0.15 (revokes), 0.29 (offending)
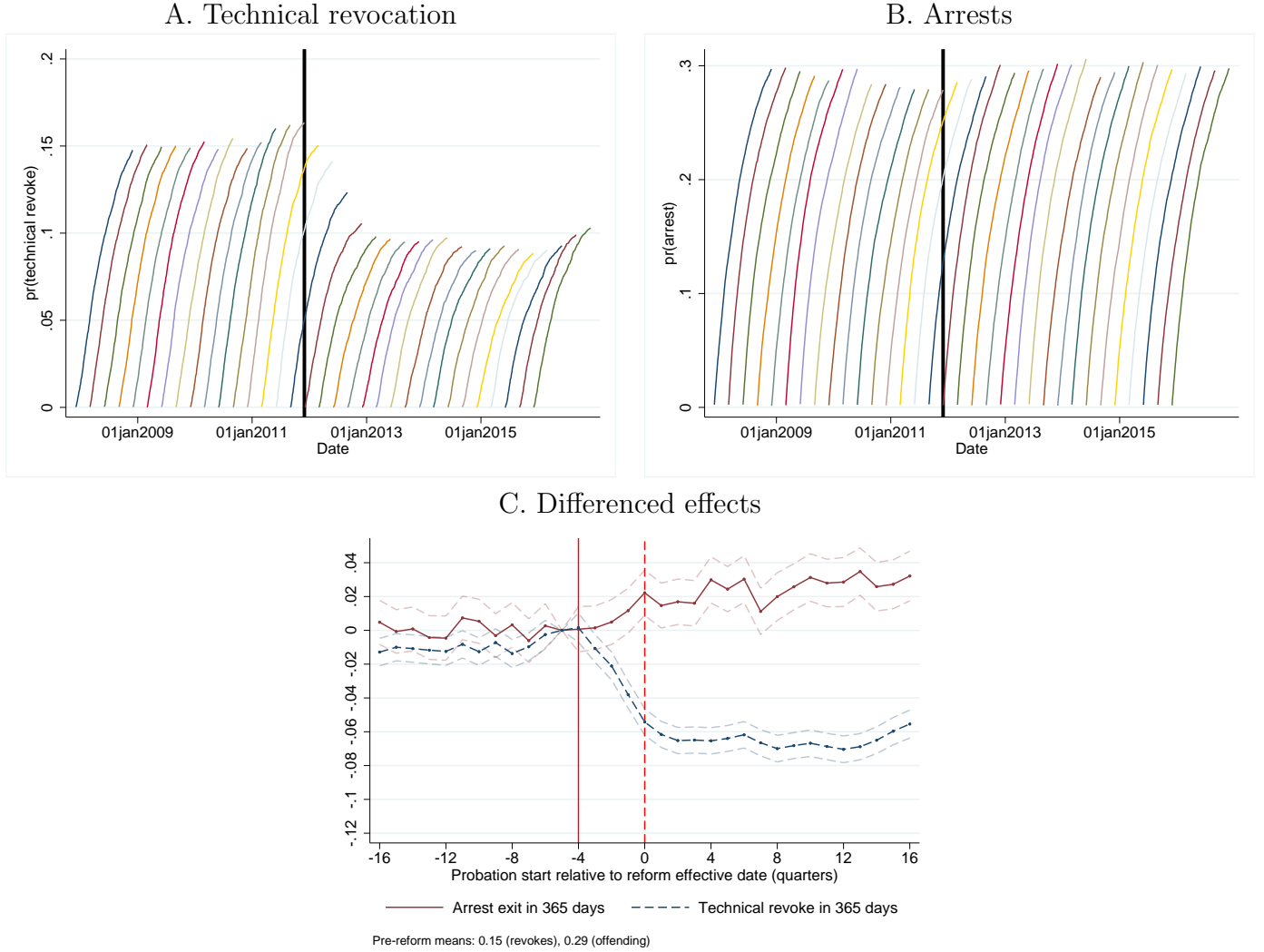
Notes: Panels A and B include all supervised probationers starting their spells within four years of the reform. Each line represents a three-month cohort of probationers who start their spells where the line intersects the x-axis. The y-axis measures the share of this cohort experiencing the relevant 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.

37

Figure 3: Effects of Reform by Race

A. Technical revocation

B. Arrests
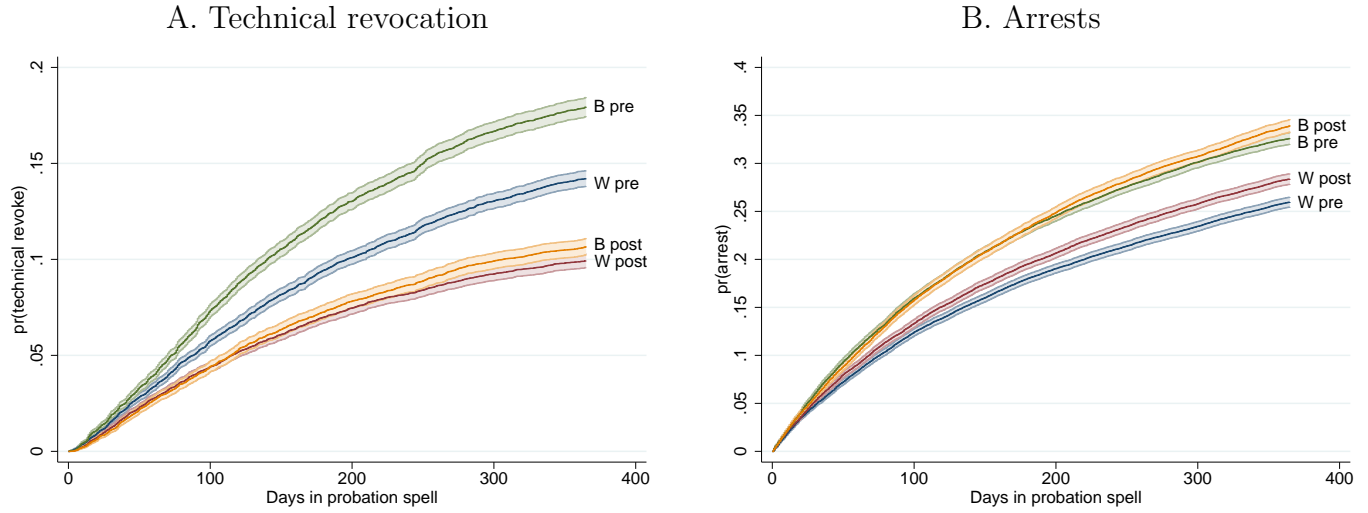
Notes: Includes all supervised probationers starting their spells either 1-3 years before (pre) or 0-2 years after the reform (post). "B" refers to black probationers, while "W" refers to non-black. The y-axis measures the share of each group experiencing the relevant outcome over the first year of their 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 usings standard errors clustered at the individual level.

Figure 4: Average Hazards for Arrest and Technical Revocation

A. Arrests

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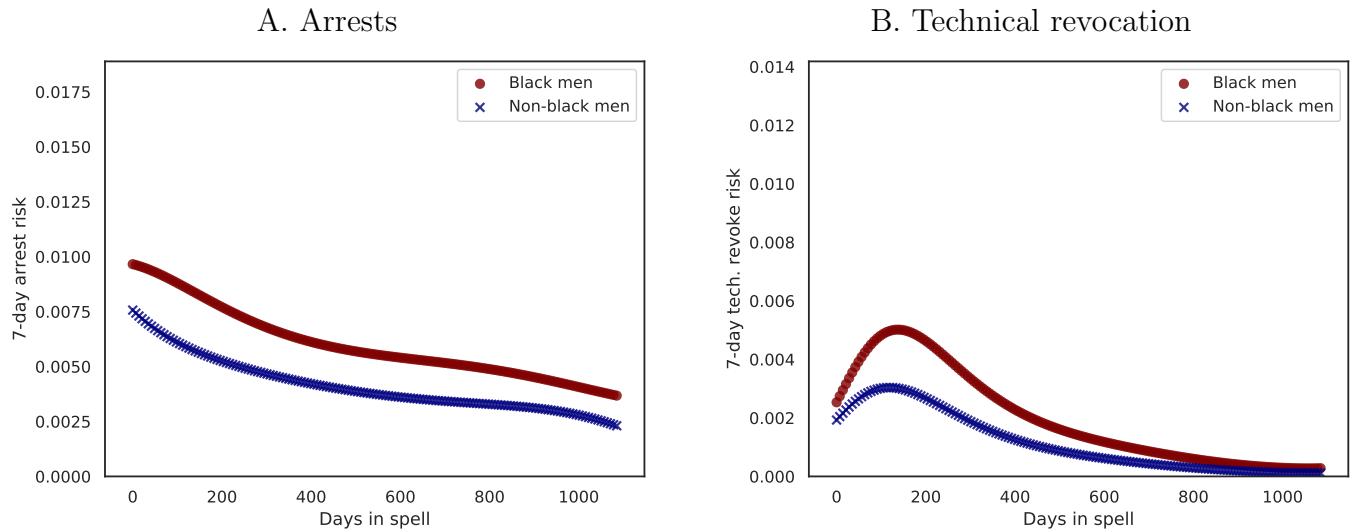
Notes: Figure plots average baseline weekly hazard rates for each outcome implied by estimates of the mixed logit competing risks model. The baseline hazard reflects the risk of each event for the *same individual* conditional on the event not happening previously. Hazards are calculated for an individual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

Figure 5: Efficiency and Equity of Technical Violation Rule Types
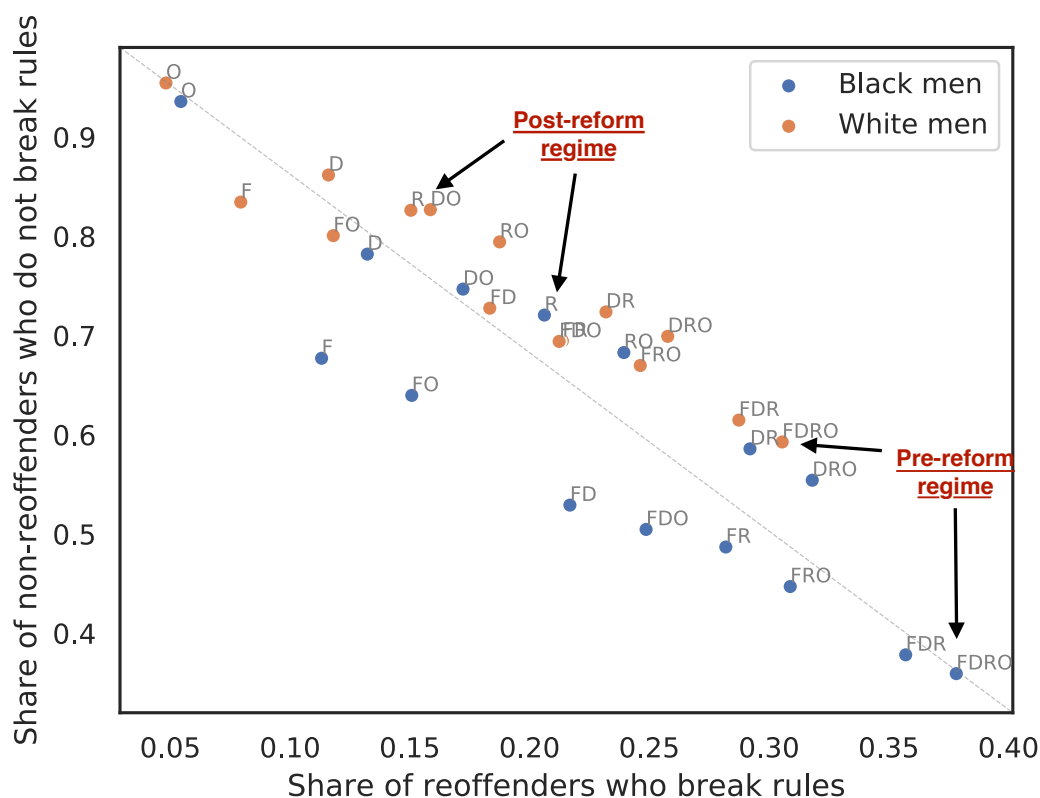
Notes: Figure plots estimates of the share of potential reoffenders over a three year period who break technical rules before they reoffend (x-axis) against the share of non-reoffenders who do not break technical rules. Estimates come from simulating the model estimated in Section 5.4 using a different set of rules (y-axis). 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 dotted gray line starts at $(1,0)$ and has a slope of -1. This line reflects what would be achieved by randomly incarcerating a fraction of probationers at the start of their spells, which naturally would catch equal shares of reoffenders and non-reoffenders.

Figure 6: Top States by Share of Prison Admissions Due to Technical Violations
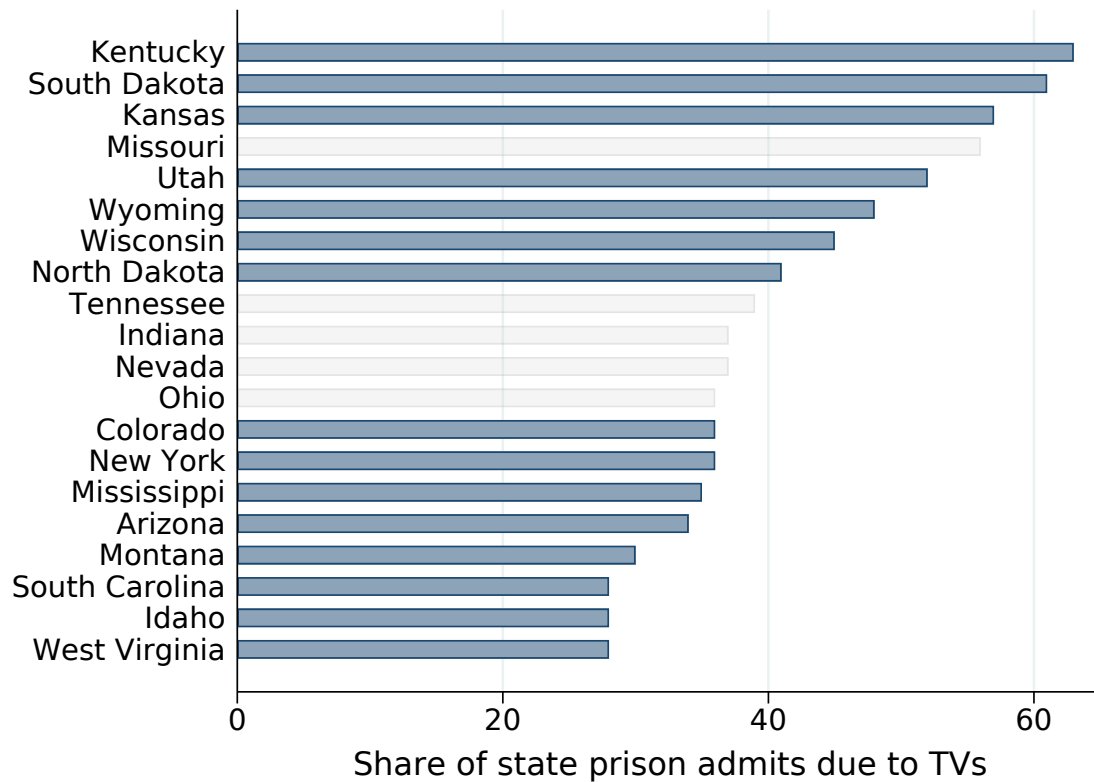
Notes: Figure plots the share of state prison admissions due to technical violations of probation and parole using data from the Council of State Governments Justice Center (CSG, 2019) for the 20 states with the highest shares. States with blue bars have no statutory limits on which technical violations can result in prison time, while states with grey bars restrict incarceration for failure to pay fees and fines when the defendant can demonstrate a financial "hardship."

Tables

Table 1: Descriptive Statistics

	Supervised (treated)			Unsupervised (control)		
	Mean	Sd.	p50	Mean	Sd.	p50
Demographics:						
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
Sentence:						
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
Criminal history:						
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
N	708623			895090		
Individuals	531099			661103		

Notes: Treated and control samples include all supervised and unsupervised probation spells beginning between 2006 and 2018, respectively. 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 2: 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
	All others	0.162	0.558

Notes: 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 3: Behavioral Responses to Reform

	Arrest		Any violation		Drug use		Fees and fines	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post reform	-0.000972	0.00133	-0.0230*	-0.0180	0.0163	0.0225	-0.0000153	0.00582
	(0.0117)	(0.0117)	(0.0101)	(0.0101)	(0.0176)	(0.0176)	(0.0118)	(0.0118)
N	152734	152734	152734	152734	152734	152734	152734	152734
Controls		Yes		Yes		Yes		Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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 are clustered at the individual level.

Table 4: Difference-in-Differences Estimates of Reform Impacts

	Technical revoke		Arrest	
A. All offenders				
	(1)	(2)	(3)	(4)
Post-reform	-0.00172***	-0.00203***	-0.00787***	-0.00699***
	(0.000274)	(0.000290)	(0.00167)	(0.00159)
Treated	0.147***	0.136***	0.0306***	-0.0156***
	(0.00105)	(0.00102)	(0.00166)	(0.00164)
Post-x-treat	-0.0546***	-0.0546***	0.0199***	0.0198***
	(0.00137)	(0.00136)	(0.00242)	(0.00233)
N	546006	546006	546006	546006
Pre-reform treated mean	.154	.154	.286	.286
Accuracy			.365 (.044)	.365 (.042)
False negative rate			.935 (.008)	.935 (.007)
False positive rate			.058 (.004)	.058 (.004)
B. Non-black offenders				
Post-reform	-0.000522	-0.000867**	-0.00688***	-0.00661***
	(0.000317)	(0.000336)	(0.00199)	(0.00190)
Treated	0.126***	0.114***	0.0442***	-0.000306
	(0.00131)	(0.00127)	(0.00208)	(0.00207)
Post-x-treat	-0.0366***	-0.0371***	0.0201***	0.0182***
	(0.00175)	(0.00174)	(0.00304)	(0.00295)
N	328784	328784	328784	328784
Pre-reform treated mean	.131	.131	.264	.264
Accuracy			.549 (.083)	.543 (.079)
False negative rate			.929 (.01)	.93 (.01)
False positive rate			.027 (.005)	.027 (.005)
C. Black offenders				
Post-reform	-0.00389***	-0.00411***	-0.0117***	-0.0111***
	(0.000509)	(0.000538)	(0.00295)	(0.00281)
Treated	0.172***	0.164***	-0.00603*	-0.0467***
	(0.00168)	(0.00168)	(0.00274)	(0.00268)
Post-x-treat	-0.0760***	-0.0756***	0.0232***	0.0237***
	(0.00217)	(0.00216)	(0.00399)	(0.00383)
N	217222	217222	217222	217222
Pre-reform treated mean	.181	.181	.314	.314
Accuracy			.305 (.052)	.306 (.049)
False negative rate			.931 (.011)	.931 (.011)
False positive rate			.095 (.007)	.094 (.007)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: 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 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 are clustered at the individual level.

Table 5: Decomposition of Racial Gaps in Technical Revocations Using One-Period Model

	Overall rates		Decomposition	
	Non-black	Black	Difference	Share of gap
Probability of technical revoke				
$\quad Pr(R_i(0) = 1)$	0.040	0.085	0.045	100.0%
Distribution of risk				
$\quad Pr(Y_i(0) = 1)$	0.314	0.377	0.063	9.7%
$\quad Pr(Y_i(0) = 0)$	0.686	0.623	-0.063	-13.4%
True positive / error rates				
$\quad Pr(R_i(0) = 1 \mid Y_i(0) = 1)$	0.071	0.069	-0.002	-1.3%
$\quad Pr(R_i(0) = 1 \mid Y_i(0) = 0)$	0.027	0.095	0.068	105.0%

Notes: 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 $R_i(1) = 0$ ($\approx 90\%$ 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 4. 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 6: Triple Difference Estimates of Differential Effect on Black Offenders

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

Standard errors in parentheses

$^*\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Notes: 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. All controls are interacted with treatment, post, and treatment times post indicators. Standard errors are clustered at the individual level.

Table 7: Cost-Benefit Analysis of Reform

	(1)	(2)	(3)	(4)	(5)	(6)
	$\triangle$ in rev. \$	$\triangle$ indir. \$	Break-even	Break-even fel.	Cost lb	Cost ub
All	-676***	246*	39,813***	100,863**	23,512	195,295
	(26)	(118)	(10,079)	(31,183)	(36,126)	(109,304)
Non-black	-450***	213	24,991*	50,576*	2,114	47,363
	(34)	(128)	(10,343)	(22,161)	(39,639)	(120,331)
Black	-957***	296	50,037**	188,899	36,439	339,574
	(40)	(224)	(17,379)	(107,553)	(62,285)	(189,895)
Non-black men	-533***	197	31,863*	55,798*	-13,146	39,561
	(43)	(164)	(13,243)	(23,950)	(43,565)	(136,574)
Black men	-1,085***	376	44,156*	149,230	38,920	340,983
	(50)	(297)	(17,615)	(87,676)	(68,152)	(206,603)

Notes: 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.

Appendix

A1   Dynamic accuracy and error rates

Technical revocation depends on the reform as $R_i(0), R_i(1)$. I modify the standard monotonicity assumption to specify that $R_i(1) \geq R_i(0) \; \forall i$. Individuals shifted by the reform from revocation in the first year of their spell to never being revoked, for example, have $R_i(0) < 365, R_i(1) = \infty$.

Potential reoffending depends on whether and when you are revoked, with $Y_i = Y_i(R_i)$. Each individual has many such potential outcomes, one for each value of R_i. Using this notation, it is straightforward to show that:

$$E[1\{Y_i = k\}1\{R_i > k\}|Z_i = 1] - E[1\{Y_i = k\}1\{R_i > k\}|Z_i = 0]$$
$$= Pr(Y_i = k, R_i > k|Z_i = 1) - Pr(Y_i = k, R_i > k|Z_i = 0)$$
$$= Pr(Y_i(R_i(1)) = k, R_i(0) > k, R_i(1) > k)$$
$$+ Pr(Y_i(R_i(1)) = k, R_i(0) < k, R_i(1) > k)$$
$$- Pr(Y_i(R_i(0)) = k, R_i(0) > k, R_i(1) > k)$$

If the first and third terms in the final equality cancel, we are left with k-specific accuracy after rescaling by the first stage $Pr(R_i(0) < k < R_i(1))$. The first assumption required to ensure these two terms do cancel incorporates the mechanical fact that if an individual is rearrested at time k, they cannot be technically revoked afterwards by definition. Hence $Y_i(k) > k$ unless $k = \infty$. Among those with $Y_i(R_i(0)) = k$, therefore, $R_i(0) = \infty$ and $R_i(1) = \infty$ by monotonicity.

The second assumption requires that $R_i(0) > k \to Y_i(R_i(1)) > k$. This condition requires that individuals who would be revoked later in their spell due to the reform do not reoffend before they would have been revoked absent the reform. It is a variation on the "no-behavioral-response" assumption imposed in the one-period model.

Under these assumptions, note that:

$$Pr(Y_i(R_i(1)) = k, R_i(0) > k, R_i(1) > k) = Pr(Y_i(\infty) = k, R_i(0) = \infty = R_i(1))$$
$$+ Pr(Y_i(R_i(1)) = k, k < R_i(0) < \infty = R_i(1))$$
$$+ Pr(Y_i(R_i(1)) = k, k < R_i(0) < \infty, R_i(1) < \infty)$$

The third term is zero by the first assumption. The second term is zero by the second assumption. Likewise,

$$Pr(Y_i(R_i(0)) = k, R_i(0) > k, R_i(1) > k) = Pr(Y_i(\infty) = k, R_i(0) = \infty = R_i(1))$$
$$+ Pr(Y_i(R_i(0)) = k, k < R_i(0) < \infty = R_i(1))$$
$$+ Pr(Y_i(R_i(0)) = k, k < R_i(0) < \infty, R_i(1) < \infty)$$

where the second and third terms are zero by the first assumption. Hence both objects reduce to $Pr(Y_i(\infty) = k, R_i(0) = \infty = R_i(1))$ and cancel. Intuitively these assumptions work by assuring that all of the observed increase in reoffending at time k due to the reform stems from individuals who would have otherwise been revoked before k and not afterwards.

Once we have obtained $Pr(Y_i(R_i(1)) = k, R_i(0) < k, R_i(1) > k)$, analogous rescalings to those

in the one-period model translate this joint probability into k-specific accuracy and error rates. Type-II error, or $Pr(R_i(0) > k | Y_i(R_i(1)) = k, R_i(1) > k)$, can be estimated using:

$$\frac{Pr(Y_i(R_i(1)) = k, R_i(1) > k) - Pr(Y_i(R_i(1)) = k, R_i(0) < k, R_i(1) > k)}{Pr(Y_i(R_i(1)) = k, R_i(1) > k)}$$

where $Pr(Y_i(R_i(1)) = k, R_i(1) > k)$ is directly observed in the population by the mean of $1\{Y_i = k\}1\{R_i > k\}$ when $Z_i = 1$.

Type-I error can still be estimated exactly as in the one-period model, but redefining the period to be k days long.

A2   Extension to accomodate CRVs

The JRA introduced the option for short confinement spells (CRVs) in response to violations. These were intended to substitute for revocations in situations where revocation was no longer permissible under the reform. Because of CRVs, observed offending post-reform might be lower than if *all* incarceration for technical violations was eliminated.

If CRVs are used exclusively as a substitute for revocation, however, the procedure in Section 3 still estimate a clear causal effect. Interviews conducted by the North Carolina Sentencing and Policy Advisor Commission in 2013 support this assumption. The commission notes that probation officers and judges did in fact use CRVs in settings when they would have revoked before the JRA, noting that in general "PPOs and judges looked for the same misbehaviors triggering a CRV as the ones they would have previously looked for to revoke probation for felons," and that "For misdemeanants, the CRV has essentially replaced revocations of probation for technical violations" (Hall et al., 2014).

To see what is identified when CRVs are used after the reform, index potential outcomes by R_i and C_i (for CRV). The assumption that CRVs are used exclusively as a substitute for revokes implies that $Pr(C_i = 1 | R_i(1) = R_i(0) = 0) = 0$. Then:

$$
\begin{aligned}
E[Y_i(1 - R_i)|Z_i = 1] - E[Y_i(1 - R_i)|Z_i = 0] &= E[Y_i(0,1)|R_i(1) < R_i(0), C_i(1) = 1]Pr(C_i(1) = 1, R_i(1) < R_i(0)) \\
&+ E[Y_i(0,0)|R_i(1) < R_i(0), C_i(1) = 0]Pr(C_i(1) = 0, R_i(1) < R_i(0)) \\
&= E[Y_i(0, C_i(1))|R_i(1) < R_i(0)]Pr(R_i(1) < R_i(0))
\end{aligned}
$$

Hence the reduced form effect of Z_i on $Y_i(1 - R_i)$ reveals a weighted average of complier outcomes subjected to CRVs and not subjected to CRVs. This reflects mean reoffending rates under the alternative policy

A3   Extension to difference-in-differences

Potential outcomes depend on Z_i (directly through time) and treatment R_i as $Y_i(R_i, Z_i)$. Let T_i indicate treatment group membership. Assume:

1. Common trends for controls and never-takers:

$$E[Y_i(0,1)|T_i = 0, Z_i = 1] - E[Y_i(0,0)|T_i = 0, Z_i = 0]$$
$$= E[Y_i(0,1)|R_i(0) = R_i(1) = 0, T_i = 1, Z_i = 1] - E[Y_i(0,0)|R_i(0) = R_i(1) = 0, T_i = 1, Z_i = 0]$$

2. Stable complier shares:

$$(R_i(1), R_i(0)) \perp\!\!\!\perp Z_i | T_i$$

Then the difference-in-differences reduced form on $Y_i(1 - R_i)$ identifies:

$$DiD(Y_i(1 - R_i)) = \underbrace{Pr(Y_i(0,1) = 1, R_i(1) = 0, R_i(0) = 1) \mid T_i = 1, Z_i = 1)}_{\text{Effect of interest}} -$$

$$\underbrace{(E[Y_i(0,1)|T_i = 0, Z_i = 1] - E[Y_i(0,0)|T_i = 0, Z_i = 0])\,(1 - Pr(R_i(1) = 0, R_i(0) = 0|T_i = 1))}_{\text{Bias term}}$$

To see this, first note that under the maintained assumptions:

$$E[Y_i(1 - R_i)|T_i = 1, Z_i = 1] = E[Y_i(0,1)|R_i(1) = 0, T_i = 1, Z_i = 1]Pr(R_i(1) = 0|T_i = 1, Z_i = 1)$$
$$= E[Y_i(0,1)|R_i(1) = 0, R_i(0) = 1, T_i = 1, Z_i = 1]Pr(R_i(1) = 0, R_i(0) = 1|T_i = 1)$$
$$+ E[Y_i(0,1)|R_i(1) = 0, R_i(0) = 0, T_i = 1, Z_i = 1]Pr(R_i(1) = 0, R_i(0) = 0|T_i = 1)$$
$$E[Y_i(1 - R_i)|T_i = 1, Z_i = 0] = E[Y_i(0,0)|R_i(1) = 0, R_i(0) = 0, T_i = 1, Z_i = 0]Pr(R_i(1) = 0, R_i(0) = 0|T_i = 1)$$

Because controls have $Pr(R_i = 0|T_i = 0) \approx 1$, we have:

$$E[Y_i(1 - R_i)|T_i = 0, Z_i = 1] - E[Y_i(1 - R_i)|T_i = 0, Z_i = 0] = E[Y_i(0,1)|T_i = 0, Z_i = 1] - E[Y_i(0,0)|T_i = 0, Z_i = 0]$$

Putting these results together shows that the difference-in-differences reduced form on $Y_i(1 - R_i)$ gives:

$$E[Y_i(0,1)|R_i(1) = 0, R_i(0) = 1, T_i = 1, Z_i = 1]Pr(R_i(1) = 0, R_i(0) = 1|T_i = 1)$$
$$+ Pr(R_i(1) = 0, R_i(0) = 0|T_i = 1)[E[Y_i(0,1)|R_i(1) = R_i(0) = 0, T_i = 1, Z_i = 1]$$
$$- E[Y_i(0,0)|R_i(1) = R_i(0) = 0, T_i = 1, Z_i = 0]] - E[Y_i(0,1)|T_i = 0, Z_i = 1] + E[Y_i(0,0)|T_i = 0, Z_i = 0]$$

Explained intuitively, if there is no time effect there is no bias, since subtracting the control groups time trend changes nothing. If there are time effects, then difference-in-differences is biased because although the full time effect is observed in the control group, the time effect in the treated group is muted by always takers and compliers. For example, in the extreme case where $Pr(R_i(0) = 1|T_i = 1) = 1$, then all members of the treated group are revoked pre-reform and $E[Y_i(1 - R_i)|T_i = 1, Z_i = 0] = 0$. The object of interest is directly identified by $E[Y_i(1 - R_i)|T_i = 1, Z_i = 1]$. Subtracting off the control groups time trend therefore introduces bias.

Empirically, the first component on the bias term is small for both race groups. In the primary estimates, for example, the time trend is: -0.00699. The share of never-takers is 84.6%. Hence

bias is roughly -0.001, or 2% of the post-x-treatment effect.

A4   Calculation of Oaxaca decomposition

I use the primary results from Table 4 to construct the one-period Oaxaca decomposition. The first row, which reports $Pr(R_i(0) = 1 | R_i(1) = 0)$ by race is -1 times the coefficient on post-x-treat, which is an estimate of $Pr(R_i(0) = 1, R_i(1) = 0)$, rescaled by the probability of being a potential complier, or $Pr(R_i(1) = 0)$. This probability is easily estimated as one minus the share of individuals revoked for technical violations in the first year of their spell in the post period (i.e., $E[R_i | Z_i = 1]$). That is, the sum of the constant, the treated indicator, and the post-x-treat indicator from Column 1.

The second row reports estimates of $Pr(Y_i = 1 | R_i(1) = 0)$. This object is estimated as the probability of offending within the first year of a probation spell after the reform, or the sum of the constant, the treated indicator, and the post-x-treat indicator from Column 3, again re-scaled by the estimate of $Pr(R_i(1) = 0)$. The third row is 1 minus the second row.

The fourth row is the coefficient on treat-x-post from Column 3 divided by the sum of the coefficients on post-x-treat, treat, and the constant from Column 3. That is, $Pr(Y_i(0) = 1, R_i(0) > R_i(1))/Pr(Y_i(0) = 1, R_i(1) = 0)$.

The fifth row is estimated by first subtracting the coefficient on post-x-treat in Column 3 from -1 times the coefficient on post-x-treat from Column 1. This object reflects $Pr(Y_i(0) = 0, R_i(0) > R_i(1))$. I then divide by $Pr(R_i(1) = 0)$ (i.e., the sum of the constant, treated indicator, and post-x-treat indicator from Column 1) minus $Pr(Y_i(0) = 1, R_i(1) = 0)$ (i.e., the sum of coefficients on post-x-treat, treat, and the constant from Column 3). This estimates $Pr(Y_i(0) = 0, R_i(1) = 0)$. The ratio gives the desired object, $Pr(R_i(0) = 1 | Y_i(0) = 0, R_i(1) = 0)$.

Calculation of the multi-period Oaxaca is analogous, except using the diff-in-diff where the outcome is $1\{Y_i = k\}1\{R_i > k\}$. The decomposition is then calculated as:

$$\underbrace{Pr(R_i(0) < 1080 | B_i = 1) - Pr(R_i(0) < 1080 | B_i = 0)}_{\text{difference in three year technical revokes}} = \tag{13}$$

$$\sum_{k=0}^{1080} \underbrace{Pr(Y_i = k | B_i = 0)}_{\text{white risk}} [\underbrace{Pr(R_i(0) < k | Y_i = k, B_i = 1) - Pr(R_i(0) < k | Y_i = k, B_i = 0)}_{\text{difference in true positive rates}}]$$

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

$$+ \underbrace{Pr(Y_i > 1080 | B_i = 0)}_{\text{white risk}} [\underbrace{Pr(R_i(0) < 1080 | Y_i > 1080, B_i = 1) - Pr(R_i(0) < 1080 | Y_i > 1080, B_i = 0)}_{\text{difference in false positive rates}}]$$

$$+ \underbrace{Pr(R_i(0) < 1080 | Y_i > 1080, B_i = 1)}_{\text{black false positive rates}} [\underbrace{Pr(Y_i > 1080 | B_i = 1) - Pr(Y_i > 1080 | B_i = 0)}_{\text{difference in risk}}]$$

where I have suppressed the conditioning $R_i(1) > k$ and let $Y_i = Y_i(R_i(1))$.

A5   Sensitivity to Exclusion Violations

I first consider the sensitivity of the single-difference estimates in Table

control CIs for 1year with covariates is -.0085 to .0001, point estimate -.0042.

A6   Details of Hazard Modeling

Formal identification results for this class of models were developed following Cox (1962) and Tsiatis (1975)'s original result that generally correlated unobserved heterogeneity across risks is not identified. Heckman and Honoré (1989) proved that when covariates are included, unobserved heterogeneity is identified with sufficient variation in X_i and under some regularity conditions. When the data contain multiple observations per person, these conditions can be relaxed substantially and no covariates are needed (see Honoré (1993) and Proposition 3 of Abbring and Van Den Berg (2003)). These results were developed for the standard continuous time proportional hazard model (i.e., $h_{is}(t) = \psi(t)exp(X'_{ist}\beta + U_i)$). The discrete-time logit specification used here can be viewed as an approximation to the discrete-time hazard yielded by such models, which takes the log-log form (i.e., $1 - exp(-exp(\theta_0(t) + X'_{ist}\beta + U_i)))$. The log-log link $ln(-ln(1-p))$ is extremely close to the logit transform $ln(p/(1-p))$ for small p.

Estimating the mixture model is difficult because the likelihood is not convex. To increase the chances that the results reflect a global optimum, I solve the model first without any unobserved heterogeneity. I then estimate the model many times using these parameters as a starting point and randomly varying the initial unobserved heterogeneity shares and locations. Informal exploration of results obtained using this method via a grid search of parameters suggest results consistently obtain a global maximum. Regardless, results change little when using a continuous heterogeneity model that is convex.

Comparing the model's cause specific hazards to Kaplan-Meier (KM) (Kaplan and Meier, 1958) estimates of the same objects, which are presented in Appendix Figure A11, further illustrates the impact of unobserved heterogeneity in this setting. The KM estimator is simply the weekly probability of failure for each cause conditional on not failing due to *any* cause previously. KM only accurately estimates hazards when there is no unobserved heterogeneity. In this case, unobserved heterogeneity and the positive correlation in risks both depresses the KM hazard estimates overall for each cause and exacerbates observed negative duration dependence, as is expected (Van Den Berg, 2001). KM estimates of arrest hazards, for example, suggest declines in risk of close to 66% for black men over the first year of a spell.

Are the model's functional form restrictions consistent with the data? I test the model's fit in multiple ways. First, Figure A10 compares the model's predicted increases in arrests as a result of the reform to difference-in-difference estimates of the reform's effects, an exercise similar in spirit to testing the fit of control function-based reproductions of non-parametric estimates of treatment effects (Kline and Walters, 2016; Rose and Shem-Tov, 2019). For each race-by-gender group, I estimate the increase in observed offending after 90, 180, 270, and 360 days using the same specification as in the difference-in-differences analysis, yielding a total of 16 points. I then simulate increases in offending in the model at each horizon and for each race-by-gender group

using the estimated offending and technical violation hazards and the effects of the reform on both. While difference-in-difference estimates are noisy, the model does a good job of capturing the basic pattern of effects.

Second, Appendix Figure A11 shows that the empirical hazards implied by the model closely match KM estimates. This is an important validation check, since it implies that the estimated distribution of unobserved heterogeneity, which is primary identified by repeated spells, generates empirical hazards that closely match patterns in the full population, which primarily includes offenders with just one spell. Appendix Figure A12 shows that model also does a good job of matching outcomes for offenders with exactly two spells as well. This plot compares model-based vs. observed joint probabilities of a given combination of outcomes (e.g., arrest or incarceration for technical violations) and timing (e.g., in the first quarter of the spell) in the first and second spell. Model predictions closely track observed probabilities, although the model may slightly underestimate the likelihood of arrest in the first quarter of both spells (the rightmost points).

Estimates of the model with continuous heterogeneity are presented in Appendix Tables A21 for men and A22 for women. Results change little, including important conclusions about state dependence over the spell and racial differences in the correlation between risks. The correlation between unobserved rearrest and incarceration for technical violations risk for black offenders is 0.2, for example, but is 65% higher for non-black offenders. The mixture model, however, generates slightly higher log likelihoods, indicating a better fit to the data.

Figure A1: Male High School Dropouts: Employment and Incarceration
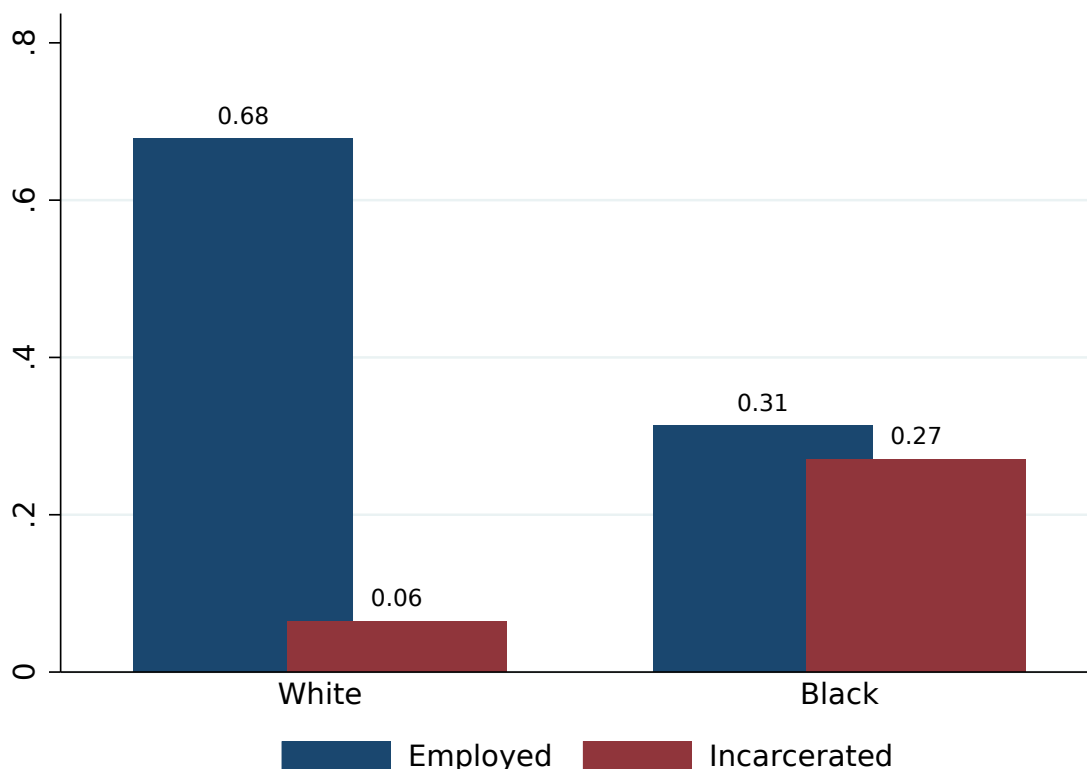
Notes: Figure constructed using the 2013-2017 5-year public use American Community Survey data (Ruggles et al., 2019). Includes White and African-American men aged 20-40 with less than 12 years of education. All estimates constructed using IPUMS person weights. Blue bars are means of an indicator for being at work at the time of enumeration. Red bars are means of an indicator for being enumerated in institutional group quarters, which includes adult correctional facilities, mental institutions, and homes for the elderly, handicapped, and poor. Breakouts for correctional facilities alone are not available in public use data, but adult correctional facilities account for 95% of the total institutional group quarters population for men 18-54 in the 2013-2017 ACS, according to Census Bureau tabulations.

Figure A2: Coefficient on Black Indicator by Detailed Violation Type
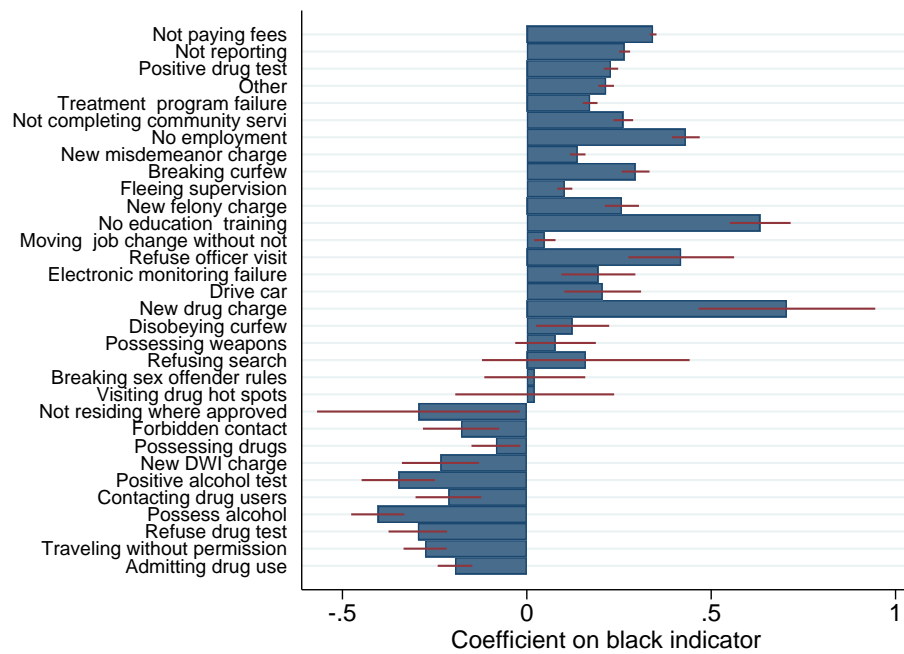
Notes: Sample and specification are the same as in Column 5 of Table A2, except the black coefficient is divided by the white mean of the dependent variable.

Figure A3: Relationship Between Racial Gaps in Technical Violations and Arrests Across North Carolina

Slope: 1.144 (0.08)
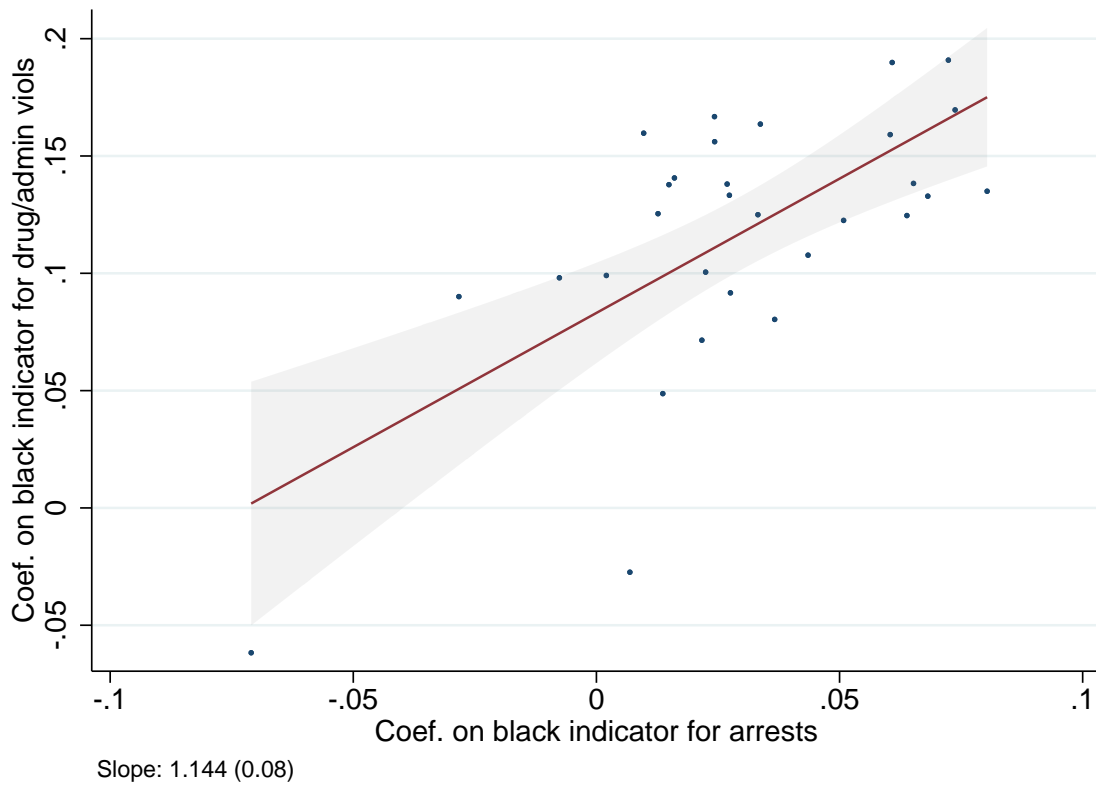
Notes: Regressions include all spells starting in 2006-2010. Each dot plots the coefficient on a black indicator from two regressions estimated separately in each of the 30 probation districts in the state. The outcome in the first regression is an indicator for any criminal arrest within three years of starting probation. The outcome in the second regression is an indicator for any drug or administrative violation in the spell. All regressions include the demographic, sentencing, and criminal history controls used in Figure 1. To avoid mechanical relationships, I randomly split the sample in half and run regressions for each outcome in separate samples, as in a split-sample IV estimate (Angrist and Krueger, 1995). The positive slope indicates that racial gaps in technical violations and racial gaps in arrest risk are positively correlated across the state, as would be expected if criminally riskier probationers incur more technical violations.

Figure A4: Illustration of Test of Behavioral Responses
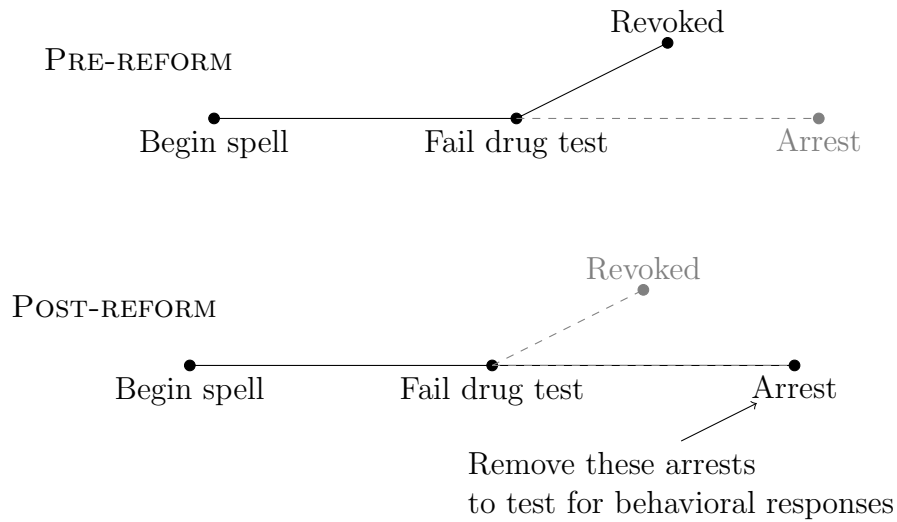
Notes: Figure illustrates the test for behavioral responses conducted in Table 3. Prior to the reform, individuals may be revoked for a rule violation such as a failed drug test. Any subsequent potential arrests would therefore be unobserved. After the reform, failed tests no longer result in incarceration, revealing previously censored arrests. By deleting all arrests that occur after technical violation, however, one can undo the impact of the reform on censoring due to technical revocation. If arrests still increase in this new measure, offenders must also respond behaviorally to the reform by increasing their criminal activity. Table 3 detects no evidence of these behavioral responses.

Figure A5: Predicted Arrest Rates Around Implementation of Reform
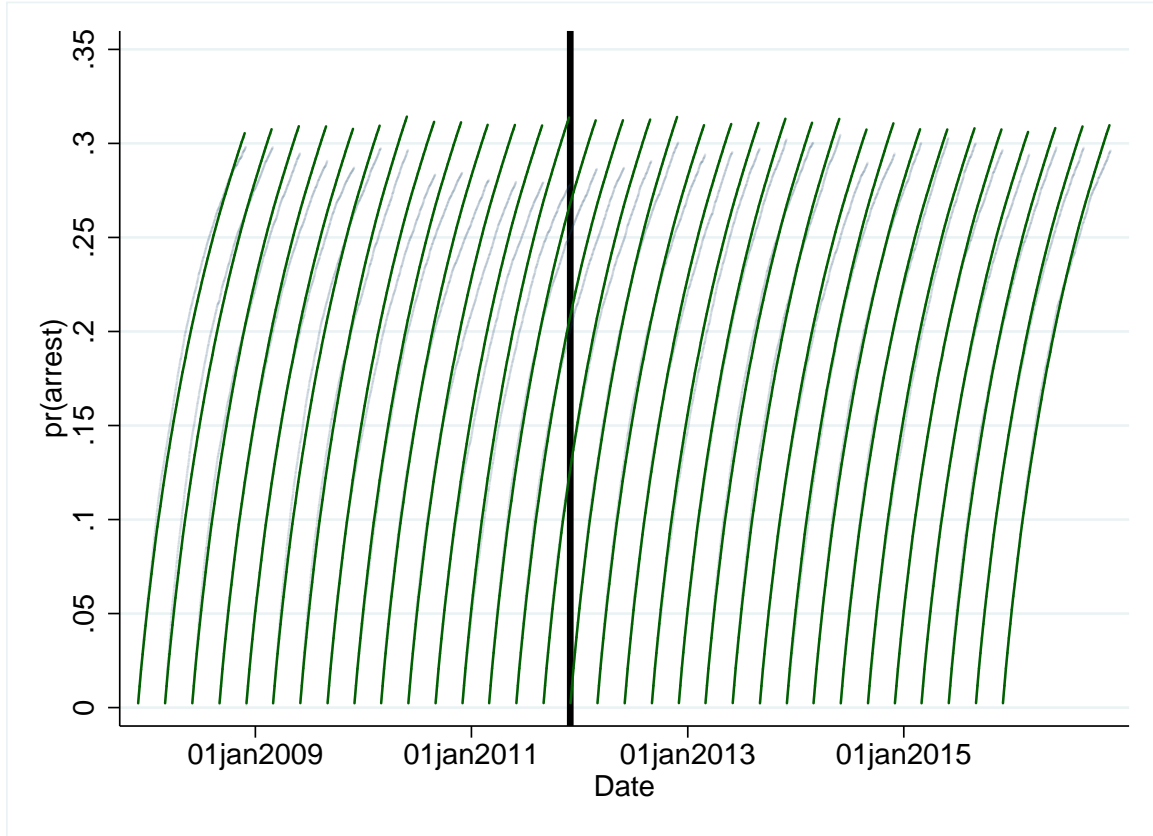
Notes: Includes all supervised probationers starting their spells within four years of the reform. Each line represents a three-month cohort of probationers who start their spells where the line intersects the x-axis. The y-axis measures the predicted share of this cohort arrested over the first year of their spells formed using linear regressions of arrest within t days on 5-year age bins interacted with race and gender, indicators for prior criminal history points and sentences to probation or incarceration, and indicators for the original arrest offense type. The regression is estimated for all $t \leq 365$ in the unsupervised (i.e., control group) probation population starting spells within 4 years of the reform. Treated (i.e., supervised) probationers' actual outcomes are reproduced in the light grey lines in the background.

Figure A6: Sample Densities Around Reform
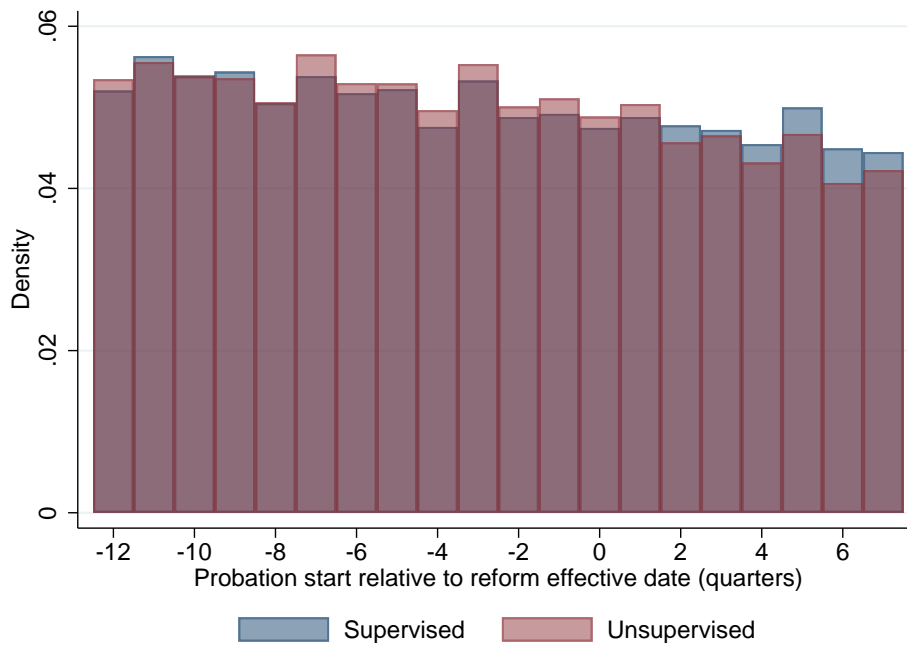
Notes: Figure plots the share of treated and untreated units in each quarter before and after the 2011 reforms for the core difference-in-differences estimates.

Figure A7: Effect of Reform on Unsupervised Probationers' Technical Revocation and Reoffending
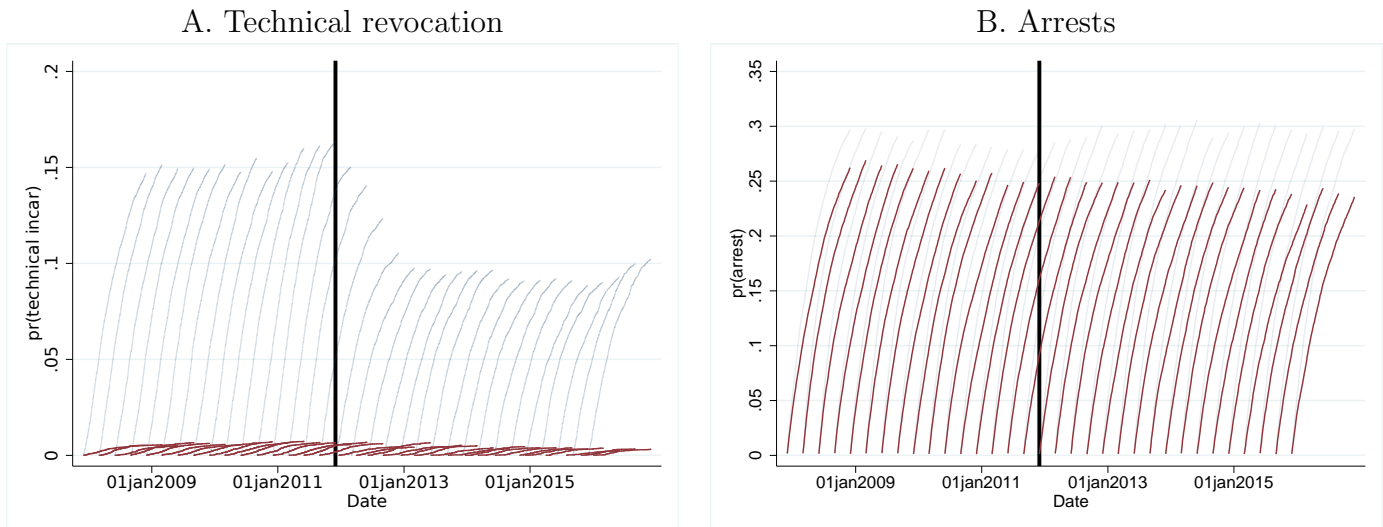
Notes: Includes all unsupervised probationers starting their spells over the included time window. Each line represents a three-month cohort of probationers starting their spells where the line intersects the x-axis. The y-axis measures the share of this cohort experiencing the relevant outcome over the following year. Technical revocation is an indicator for having probation revoked with no intervening arrest. Arrest is an indicator for being arrested before being revoked. Treated (i.e., supervised) probationers' outcomes are reproduced in the light grey lines in the background.

Figure A8: Dynamic Estimates of Error Rates
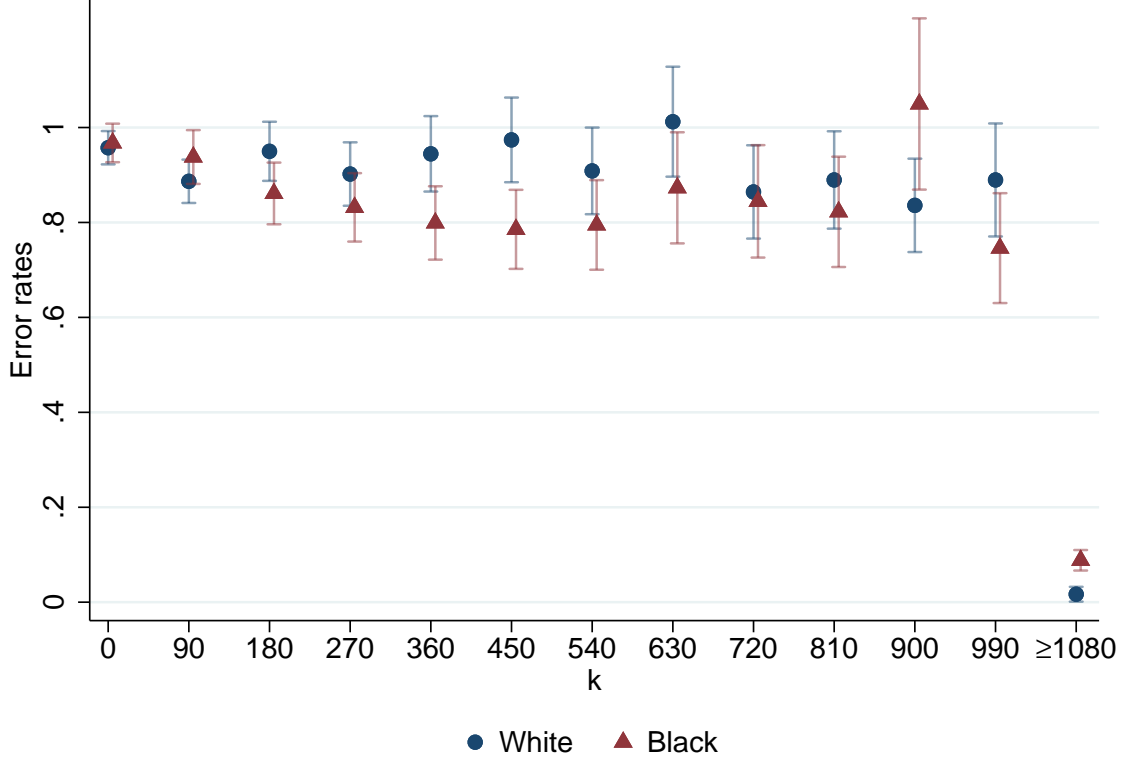
Notes: Figure plots estimates and 95% confidence intervals of type-II error rates, i.e., $Pr(R_{ik}(0) = 0 | Y_{ik}(0) = 1, R_{ik}(1) = 1)$, by race. These error rates reflect the probability probationers who would otherwise be rearrested at time k were *not* revoked for violating technical rules affected by the reform. Larger values for white offenders indicate that rules catch a larger fraction white of potential reoffenders at each horizon k. The final point, above ≥ 1080, measures the share of probationers who would not be rearrested within 1080 days revoked for violating technical rules, or type-I error at a three-year horizon. Error rates are estimated using estimates of the core difference-in-differences specification in Table 4. The outcomes for each k are Y_{ik}, an indicator for being rearrested within k and $k + 89$ days of probation start without any intervening technical revocation, and R_{ik}, an indicator for being revoked for rule violations before time k. Error rates are calculated as described in Section 3.2. Spells starting pre-reform with sentenced lengths that imply finishing post reform are dropped, since these spells are only partially affected.

Figure A9: Targeting Disparities in the Competing Risks Model
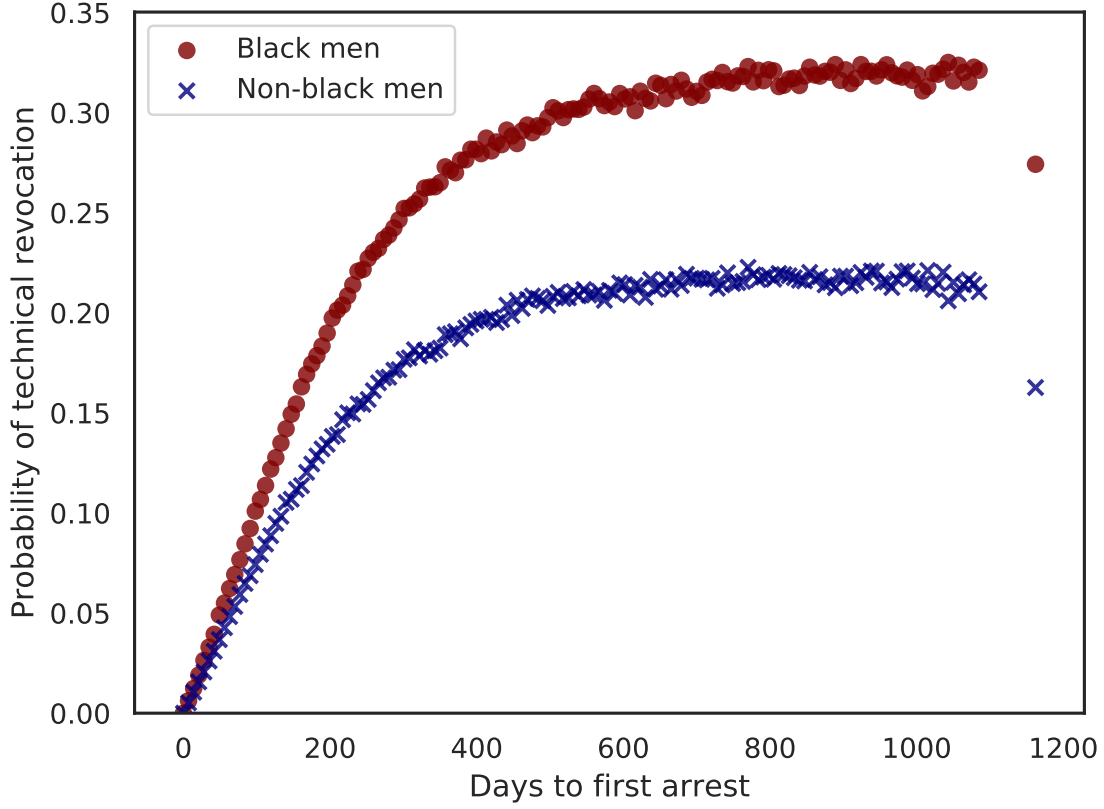
Notes: Figure plots estimates of $Pr(R_i^* < Y_i^*|Y_i^* = k)$, or the likelihood of technical revocation before time k among probationers who would be otherwise be rearrested at time k, from simulating outcomes in the competing risks model. Simulations use the pre-reform empirical distribution of covariates for each race-gender group and the estimated race-gender specific distributions of unobserved heterogeneity. $Pr(R_i^* < Y_i^*|Y_i^* = k)$ is the share of observations across simulations who have reoffending failure times equal to k but technical incarceration failure times $< k$. Higher values for black probationers indicate that among probationers who would otherwise be rearrested at the same time, technical rules target black probationers more aggressively. The final dots at the right of the graph plot the probability of technical revocation failure times ≤ 1080 conditional on having arrest failure times > 1080 (and possibly infinite).

Figure A10: Model-based Replication of Difference-in-Difference Estimates
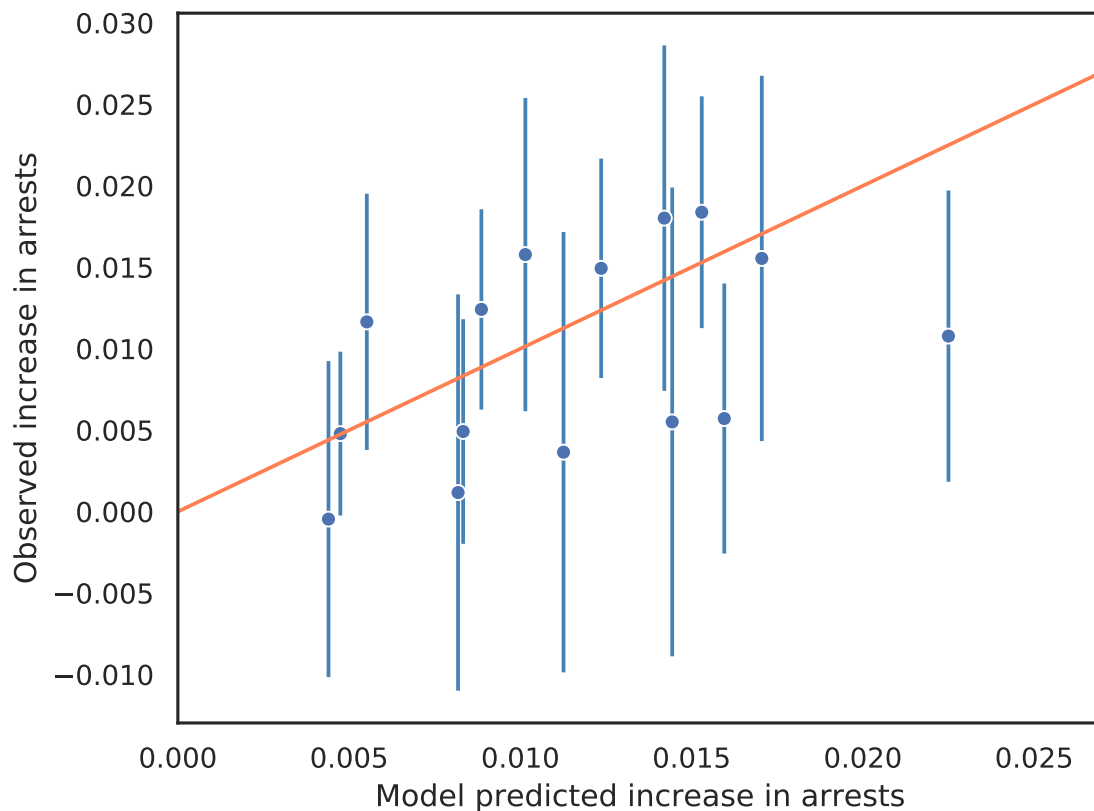
Notes: Figure compares difference-in-difference estimates of increases in observed arrests at 90, 180, 270, and 360 days for each race-by-gender group to the competing risk model's prediction of the same object. Vertical lines reflect 95% confidence intervals for the diff-in-dif estimates, while the orange line lies on a 45 degree angle. The diff-in-diff estimates are constructed using the sample sample and specification as in the reduced-form analysis and with no covariates included. Model predictions come from simulating observed arrests at each horizon with and without the "post-reform" coefficients turned on. Covariates are fixed at the empirical distribution in the pre-reform period.

Figure A11: Competing Risks Model Fit to Kaplan-Meier Estimates of Hazards

A. Black men

B. Non-black men
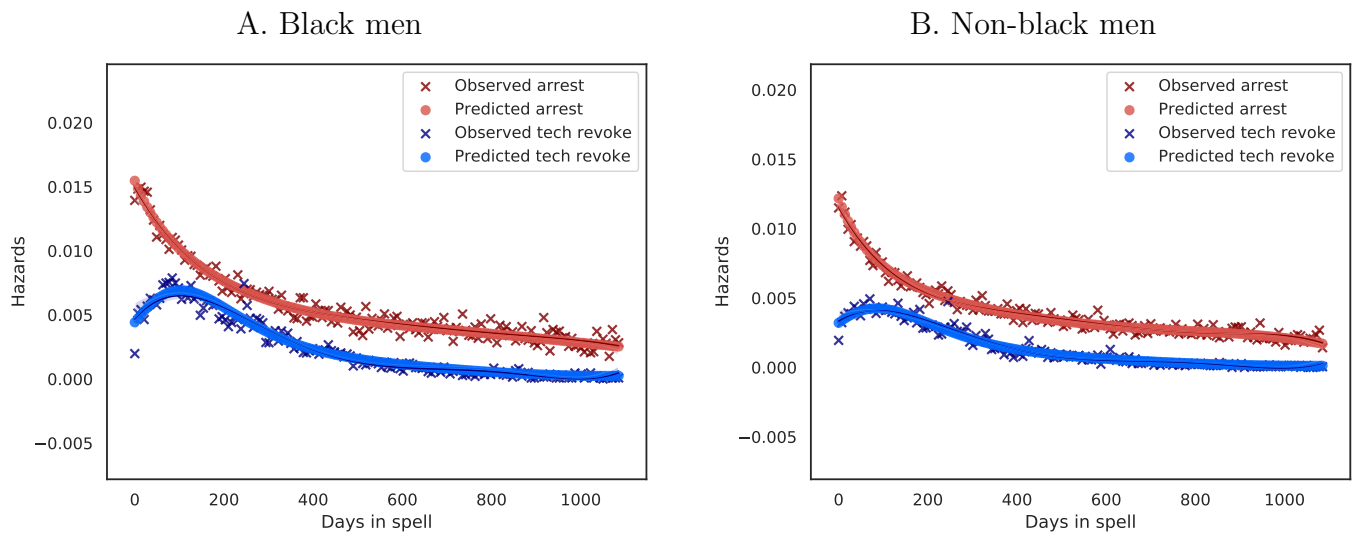
Notes: Figure plots Kaplan-Meier estimates of the cause-specific hazard for spells beginning three to one years before the reform and model simulations of the same object. The Kaplan-Meier estimator in this context is simply the weekly probability of arrest or technical incarceration conditional on neither event happening previously. Model based estimates are simulations of the sample probabilities.

Figure A12: Competing Risks Model Fit to Joint Distribution of Exits Across Repeated Spells
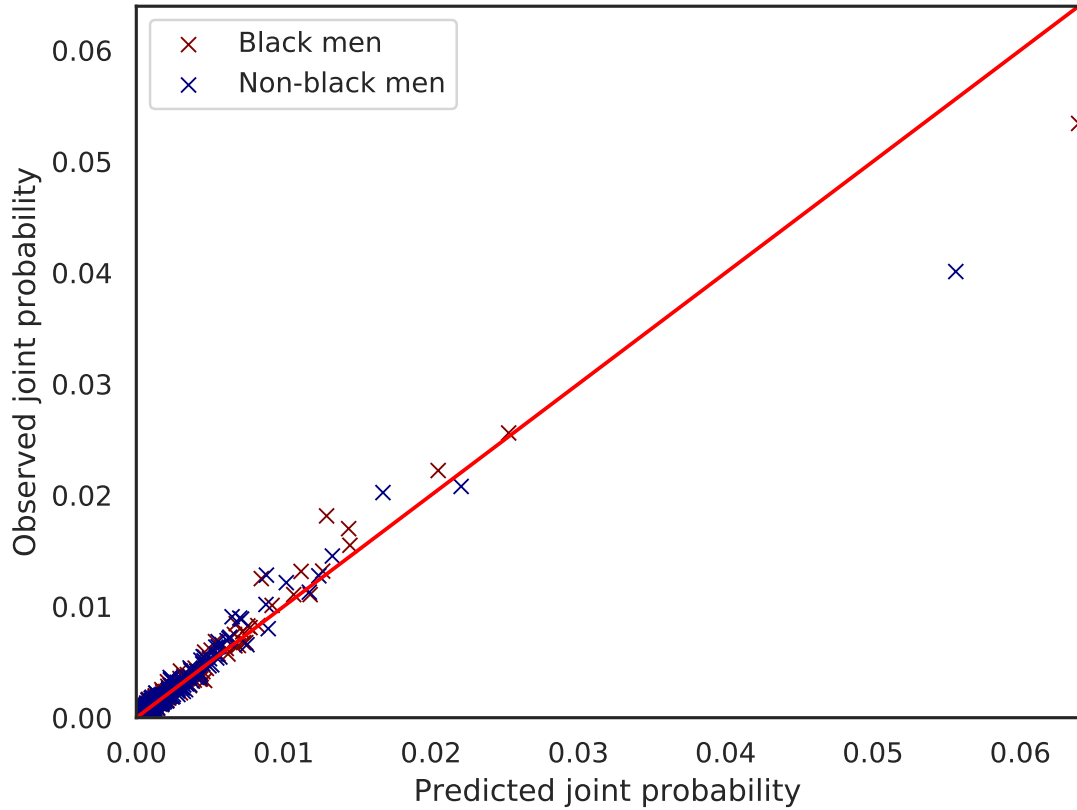
Notes: Figure plots the observed vs. predicted probabilities of failure types and times for black and non-black probationers with two probation spells. Each point in the figure is a separate failure combination across the two spells, with failure times grouped at the quarterly level. The rightmost points, for example, are the joint probabilities of being arrested in the first quarter of both spells. Other dots reflect the probability of arrest in the first quarter of the first spell, and technical incarceration of the first quarter of the second, etc. Failure times up to 12 quarters are included, yielding 12·12 combinations of possible failure times across the spells, and 4 combinations of failure types (e.g., arrest arrest, arrest tech incar, etc.), and therefore 576 points per group.

Figure A13: Impact of Reform on Baseline Hazards in Competing Risks Model

A. Arrests

B. Incarceration for technical violations
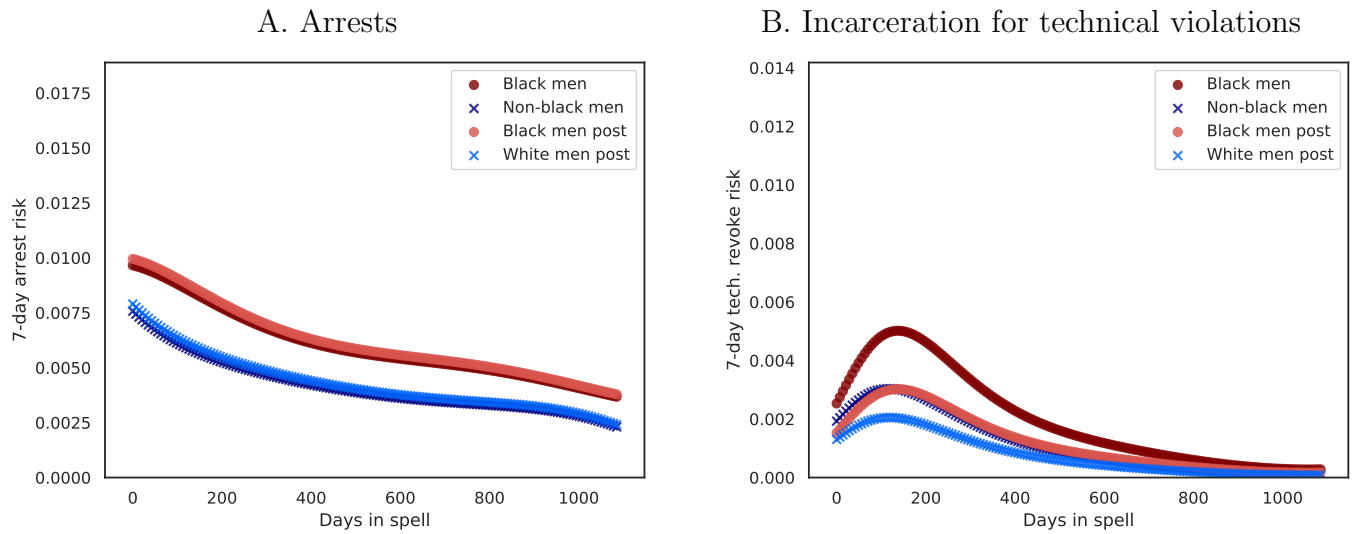
Notes: Figure plots arrest hazards for offenders with average values of the covariates implied by estimates of the competing risks model. Hazards are averaged over the distribution of unobserved heterogeneity using estimates from finite mixture version of the model estiamted with four types.

Figure A14: Targeting Bias in the Competing Risks Model Based on Unobserved Heterogeneity Only
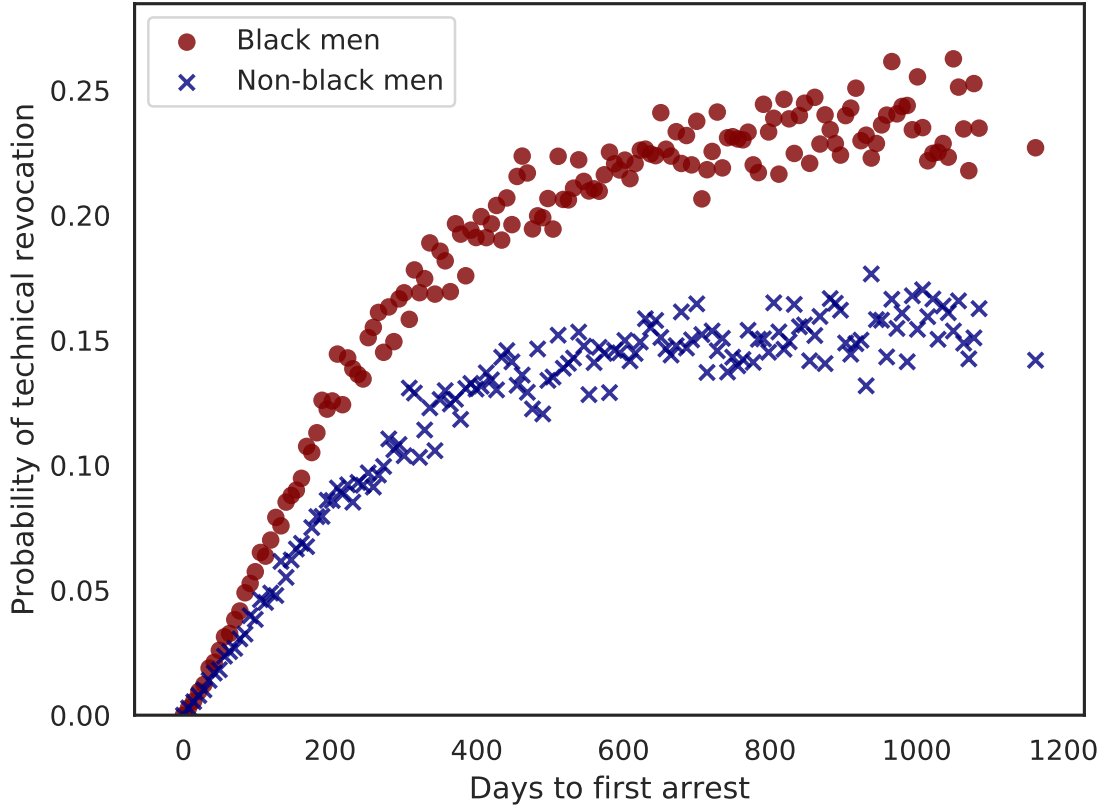
Notes: Figure plots estimates of $Pr(R_i^* < Y_i^*|Y_i^* = k)$, or the likelihood of technical revocation before time k among probationers who would be otherwise be rearrested at time k, from simulating outcomes in the competing risks model. Observables are held constant at their mean levels for men in the sample and simulations use the estimated race-gender specific distributions of unobserved heterogeneity. $Pr(R_i^* < Y_i^*|Y_i^* = k)$ is the share of observations across simulations who have reoffending failure times equal to k but technical incarceration failure times $< k$. Higher values for black probationers indicate that among probationers who would otherwise be rearrested at the same time, technical rules target black probationers more aggressively. The final dots at the right of the graph plot the probability of technical revocation failure times ≤ 1080 conditional on having arrest failure times > 1080 (and possibly infinite).

Figure A15: Average Risks for Multiple Violation Outcomes
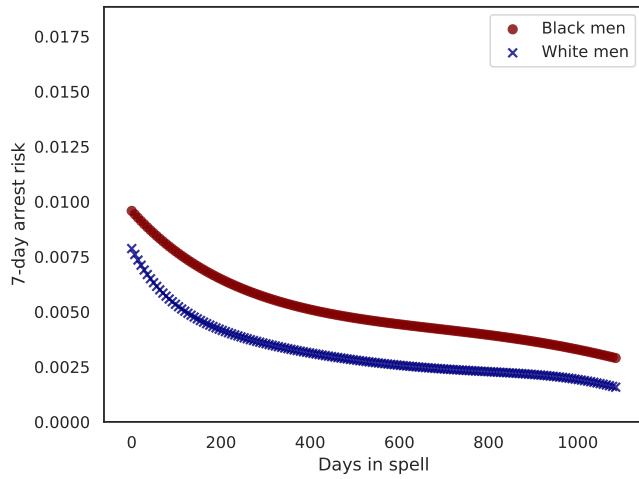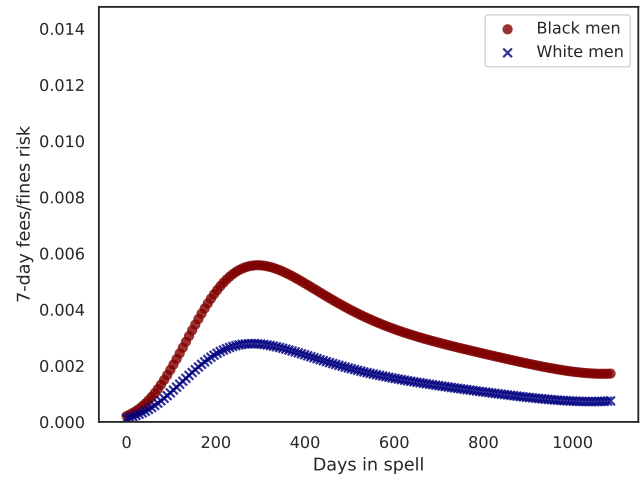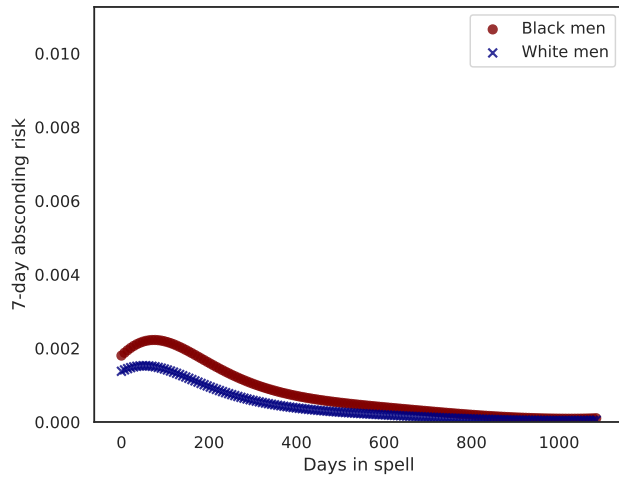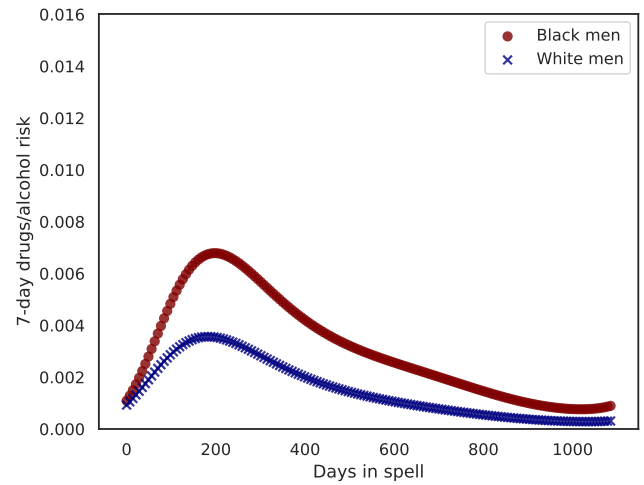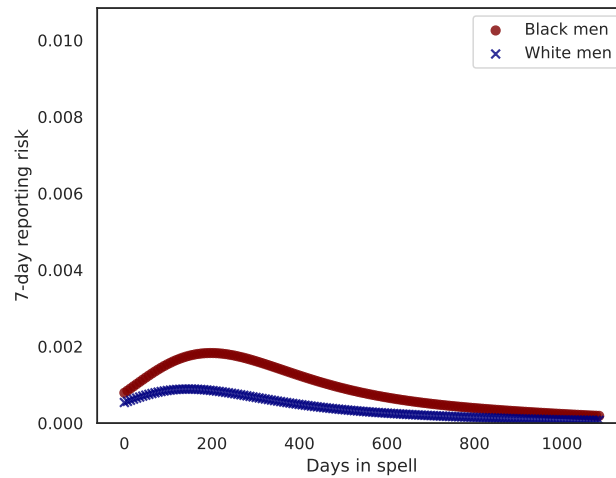
Notes: Figure plots baseline risks of committing each violation type implied by the multi-outcome competing risks model. See text for details on sample and specification of unobserved heterogeneity used in estimation. Mean weekly risks are similar but not identical to the baseline hazard, since the partial effects of unobserved heterogeneity on the hazard depend on baseline levels in the logit formulation.
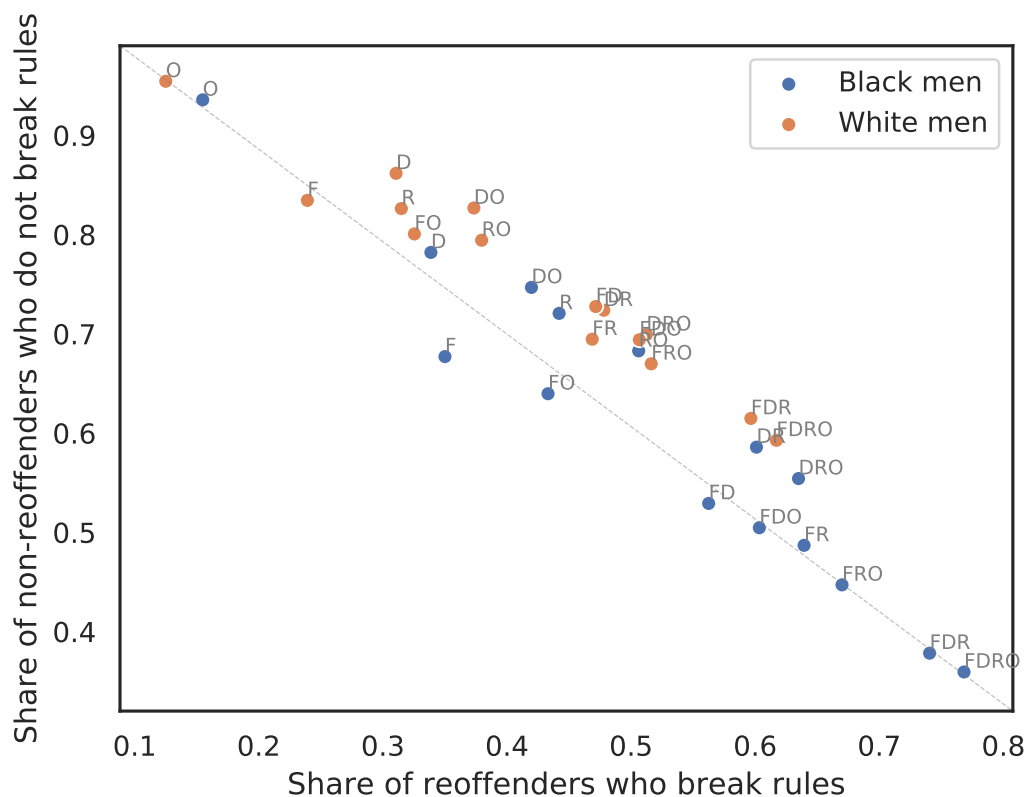
Figure A16: Efficiency and Equity of Technical Violation Rule Types Eliminating Impact of Violation Timing

Notes: Figure plots estimates of the share of potential reoffenders over a three year period who would break technical rules at any point in their spell if their arrest was ignored (x-axis) against the share of non-reoffenders who do not break technical rules. Estimates come from simulating the model estimated in Section 5.4 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 dotted grey-line starts at $(1,0)$ and has a slope of -1. This line reflects what would be achieved by randomly incarcerating a fraction of probationers at the start of their spells, which naturally would catch equal shares of re-offenders and non-reoffenders.

Table A1: Violation Categorization

Violation type	Violation	Share of category
Absconding	-	1
Drug related	Positive drug test	0.526
	Treatment / program failure	0.295
	Admitting drug use	0.071
	Possessing drugs	0.036
	Contacting drug users	0.022
New criminal offense	New misdemeanor charge	0.716
	New felony charge	0.263
	New DWI charge	0.013
	New drug charge	0.007
Technical	Not paying fees	0.427
	Not reporting	0.202
	Other	0.099
	Moving / job change without notifying	0.058
	Breaking curfew	0.055
	Not completing community service	0.047
	No employment	0.043
	No education / training	0.012
	Traveling without permission	0.011

Notes: Includes all treated observations starting probation in 2006-2010.

Table A2: Effect of Race on Administrative Violations

	Outcome: Administrative violation in spell					
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.174***	0.190***	0.177***	0.145***	0.137***	0.101***
	(0.00172)	(0.00184)	(0.00185)	(0.00183)	(0.00195)	(0.00371)
N	315167	315167	315167	315167	315167	89122
R-squared	0.0309	0.0473	0.0697	0.114	0.128	0.107
Dep. var white mean	0.512	0.512	0.512	0.512	0.512	0.512
Demographic controls		Yes	Yes	Yes	Yes	Yes
Sentence controls			Yes	Yes	Yes	Yes
Criminal history controls				Yes	Yes	Yes
Zip code FE					Yes	Yes
Test score controls						Yes
Logit coefficient	0.734	0.817	0.779	0.665		
Logit AME	0.172	0.188	0.175	0.142		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Regressions include all spells beginning in 2006-2010. Demographic controls include gender, 20 quantiles of age, and probation district fixed effects. Sentence controls include fixed effects for the offense class of the focal conviction and a linear control for the length of the supervision spell. Criminal history controls include fixed effects for criminal history points and previous sentences to supervised probation or incarceration. Zip code FE are fixed effects for zip code at the time of initial arrest. Test score controls include the latest math and reading standardized test scores (normalized to have mean 0 and standard deviation 1 in the full population) observed from grades 3 to 8. Logit coefficient and AME are the coefficient and average marginal effects from logit estimations of the same specification. These are omitted for the last two columns where the number of fixed effects is high. Stanard errors are clustered at the individual level.

Table A3: Effect of Race on Drug Violations

	Outcome: Drug violation in spell					
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.0608***	0.0677***	0.0653***	0.0448***	0.0423***	0.0212***
	(0.00162)	(0.00171)	(0.00173)	(0.00173)	(0.00184)	(0.00388)
N	315167	315167	315167	315167	315167	89122
R-squared	0.00450	0.0241	0.0396	0.0614	0.0723	0.0695
Dep. var white mean	0.257	0.257	0.257	0.257	0.257	0.257
Demographic controls		Yes	Yes	Yes	Yes	Yes
Sentence controls			Yes	Yes	Yes	Yes
Criminal history controls				Yes	Yes	Yes
Zip code FE					Yes	Yes
Test score controls						Yes
Logit coefficient	0.298	0.340	0.331	0.233		
Logit AME	0.0603	0.0675	0.0646	0.0444		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: See notes to Table A2.

Table A4: Effect of Race on Absconding Violations

	Outcome: Absconded in spell					
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.0422***	0.0503***	0.0427***	0.0232***	0.0151***	0.0132***
	(0.00135)	(0.00143)	(0.00144)	(0.00144)	(0.00153)	(0.00317)
N	315167	315167	315167	315167	315167	89122
R-squared	0.00318	0.0176	0.0279	0.0555	0.0683	0.0725
Dep. var white mean	0.147	0.147	0.147	0.147	0.147	0.147
Demographic controls		Yes	Yes	Yes	Yes	Yes
Sentence controls			Yes	Yes	Yes	Yes
Criminal history controls				Yes	Yes	Yes
Zip code FE					Yes	Yes
Test score controls						Yes
Logit coefficient	0.303	0.367	0.310	0.181		
Logit AME	0.0418	0.0498	0.0417	0.0235		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: See notes to Table A2

Table A5: Effect of Race on Revocations

	Outcome: Revoked					
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.104***	0.118***	0.105***	0.0672***	0.0599***	0.0518***
	(0.00170)	(0.00179)	(0.00181)	(0.00177)	(0.00188)	(0.00390)
N	315167	315167	315167	315167	315167	89122
R-squared	0.0118	0.0397	0.0595	0.121	0.133	0.127
Dep. var white mean	0.296	0.296	0.296	0.296	0.296	0.296
Demographic controls		Yes	Yes	Yes	Yes	Yes
Sentence controls			Yes	Yes	Yes	Yes
Criminal history controls				Yes	Yes	Yes
Zip code FE					Yes	Yes
Test score controls						Yes
Logit coefficient	0.459	0.543	0.488	0.339		
Logit AME	0.102	0.117	0.103	0.0669		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: See notes to Table A2

Table A6: Effect of Race on Technical Revocations

	Outcome: Technical revocation in spell					
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.0627***	0.0710***	0.0649***	0.0485***	0.0418***	0.0316***
	(0.00139)	(0.00147)	(0.00150)	(0.00150)	(0.00159)	(0.00334)
N	315167	315167	315167	315167	315167	89122
R-squared	0.00664	0.0153	0.0219	0.0404	0.0503	0.0484
Dep. var white mean	0.150	0.150	0.150	0.150	0.150	0.150
Demographic controls		Yes	Yes	Yes	Yes	Yes
Sentence controls			Yes	Yes	Yes	Yes
Criminal history controls				Yes	Yes	Yes
Zip code FE					Yes	Yes
Test score controls						Yes
Logit coefficient	0.426	0.488	0.448	0.345		
Logit AME	0.0619	0.0704	0.0641	0.0485		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: See notes to Table A2

Table A7: Effect of Race on Criminal Arrests

	Outcome: Arrested in spell					
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.0627***	0.0690***	0.0562***	0.0284***	0.0300***	0.0310***
	(0.00172)	(0.00182)	(0.00184)	(0.00183)	(0.00194)	(0.00402)
N	315167	315167	315167	315167	315167	89122
R-squared	0.00423	0.0284	0.0453	0.0788	0.0893	0.0742
Dep. var white mean	0.330	0.330	0.330	0.330	0.330	0.330
Demographic controls		Yes	Yes	Yes	Yes	Yes
Sentence controls			Yes	Yes	Yes	Yes
Criminal history controls				Yes	Yes	Yes
Zip code FE					Yes	Yes
Test score controls						Yes
Logit coefficient	0.272	0.308	0.253	0.133		
Logit AME	0.0623	0.0688	0.0555	0.0282		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: See notes to Table A2

Table A8: Effect of Race on Revocation Conditional on Violation

	Outcome: Revoked (conditional on violation)				
	(1)	(2)	(3)	(4)	(5)
Black	-0.00444*	0.00829***	0.00304	-0.0112***	0.00241
	(0.00180)	(0.00193)	(0.00195)	(0.00193)	(0.00208)
N	296369	296369	296369	296369	296369
R-squared	0.0000205	0.0225	0.0308	0.0562	0.406
Dep. var white mean	0.401	0.401	0.401	0.401	0.401
Demographic controls		Yes	Yes	Yes	Yes
Sentence controls			Yes	Yes	Yes
Criminal history controls				Yes	Yes
Violations FE					Yes
Logit coefficient	-0.0185	0.0358	0.0139	-0.0479	
Logit AME	-0.00444	0.00838	0.00323	-0.0108	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Includes all violation hearings for spells beginning in 2006-2010. Controls are as defined in Table A2, except for violations FE, which are fixed effects for the unique violations categories disposed at the hearing. Logit coefficient and AME are the coefficient and average marginal effects from logit estimations of the same specification. These are omitted for specifications where the number of fixed effects is high. Standard errors are clustered at the individual level.

Table A9: Officer-Offender Race Match Effect in Violations

	Outcome: Any outcome in spell							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Adm	Adm	Drug	Drug	Rev.	Rev.	Tech rev.	Tech rev.
Black	0.092***	0.091***	0.026***	0.024***	0.040***	0.041***	0.031***	0.033***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Black x black off		0.0028		0.0075*		-0.0041		-0.0044
		(0.004)		(0.003)		(0.003)		(0.003)
N	306418	306418	306418	306418	306418	306418	306418	306418
W mean	0.37	0.37	0.18	0.18	0.21	0.21	0.12	0.12
Demo	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sent	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crim hist	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Off FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Includes all spells starting in 2006-2010. Officer race is coded using the race of the first officer assigned in the spell. Controls are as defined in Table A2. Outcomes are an indicator for the listed event happening within the first year of a spell. Standard errors are clustered at the individual level.

Table A10: Difference-in-Differences Tests of Covariate Balance

	All					
	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{Arrest}$	Age	Male	Prob length	Prior sent.	Prior points
	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{Arrest}$	Age	Male	Prob length	Prior sent.	Prior points
Post-reform	0.000573	0.838***	-0.0121***	-0.411***	0.117***	0.116***
	(0.000466)	(0.0404)	(0.00167)	(0.0269)	(0.00587)	(0.00832)
Treated	0.0494***	-0.573***	0.0132***	5.466***	0.858***	0.605***
	(0.000476)	(0.0402)	(0.00162)	(0.0333)	(0.00732)	(0.00885)
Post-x-treat	-0.00107	-0.471***	-0.00260	0.0676	0.00256	-0.0414**
	(0.000680)	(0.0582)	(0.00237)	(0.0473)	(0.0107)	(0.0129)
N	546006	546006	546006	546006	546006	546006
Pre-reform treated mean	.31	31.691	.747	18.838	1.76	1.704
	Black					
Post-reform	-0.000216	0.543***	-0.0000566	-0.333***	0.169***	0.113***
	(0.000848)	(0.0683)	(0.00277)	(0.0454)	(0.0107)	(0.0161)
Treated	0.0401***	-1.127***	0.0382***	5.957***	0.781***	0.420***
	(0.000805)	(0.0649)	(0.00255)	(0.0507)	(0.0120)	(0.0154)
Post-x-treat	-0.00142	-0.242**	-0.00406	-0.297***	-0.0358*	-0.0497*
	(0.00114)	(0.0933)	(0.00368)	(0.0723)	(0.0176)	(0.0225)
N	217222	217222	217222	217222	217222	217222
Pre-reform treated mean	.354	31.33	.768	18.81	1.976	2.027
	Non-Black					
Post-reform	0.00000281	1.004***	-0.0189***	-0.446***	0.0819***	0.109***
	(0.000495)	(0.0501)	(0.00209)	(0.0333)	(0.00676)	(0.00901)
Treated	0.0434***	-0.172***	-0.00613**	5.199***	0.847***	0.626***
	(0.000536)	(0.0516)	(0.00213)	(0.0446)	(0.00919)	(0.0105)
Post-x-treat	0.00142	-0.593***	-0.00356	0.329***	0.0326*	-0.0169
	(0.000773)	(0.0749)	(0.00312)	(0.0631)	(0.0134)	(0.0153)
N	328784	328784	328784	328784	328784	328784
Pre-reform treated mean	.274	31.983	.729	18.861	1.585	1.442

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The outcome is listed in the column header. $\hat{Arrest}$ is predicted arrest rates from a linear regression of an indicator for arrest within 1 year of starting probation on gender, race, indicators for five year age bins, the interactions of these terms, fixed effects for prior record points, and fixed effects for prior sentences to DPS supervision or incarceration using all observations starting probation more than 3 years before the reform. "Post reform" is a indicator for starting probation after Dec. 1, 2011. Includes all spells starting 0-2 years after the reform or 1-3 years before. Standard errors are clustered by individual.

Table A11: Single-Difference Estimates of Effect of Reform

	Black		Non-black	
	(1)	(2)	(3)	(4)
	Revoke	Arrest	Revoke	Arrest
Post-reform	-0.0748***	0.0123**	-0.0433***	0.0183***
	(0.00298)	(0.00391)	(0.00244)	(0.00336)
N	52397	52397	65335	65335
Pre-reform treated mean	.181	.31	.136	.257
Demographic controls	Yes	Yes	Yes	Yes
Criminal history FE	Yes	Yes	Yes	Yes
Accuracy		.172 (.05)		.454 (.07)
False negative		.96 (.012)		.931 (.012)
False positive		.108 (.007)		.037 (.006)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Includes all treated probation spells beginning 1-2 years before the reform and 0-1 years afterwards. Post is indicator for starting probation after December 1, 2011, the date JRA reforms took effect. Demographic controls include five-year age bins and indicators for race and gender. Criminal history controls include fixed effects criminal history points and prior sentences to supervised probation or incarceration. Standard errors are clustered by individual.

Table A12: Effect of Reform on Felony Technical Incarceration

	Black				Non-black			
	(1) Tech Inc	(2) Any Tech Inc	(3) 2SLS Inc	(4) 2SLS Any	(5) Tech Inc	(6) Any Tech Inc	(7) 2SLS Inc	(8) 2SLS Any
Post-reform	-1.412***	-0.00771***	-0.655	-0.00553***	-0.440	-0.00386***	-0.341	-0.00359***
	(0.401)	(0.000763)	(0.444)	(0.000874)	(0.310)	(0.000446)	(0.319)	(0.000467)
Treated	16.30***	0.0834***	0.935	0.0392***	17.79***	0.0903***	5.462**	0.0568***
	(0.836)	(0.00265)	(1.216)	(0.00332)	(0.865)	(0.00261)	(1.855)	(0.00445)
Post-x-treat	-10.28***	-0.0296***			-6.518***	-0.0177***		
	(1.175)	(0.00351)			(1.282)	(0.00359)		
Revoked			204.1***	0.587***			172.6***	0.469***
			(21.73)	(0.0554)			(31.86)	(0.0776)
N	140935	140936	140935	140936	227556	227557	227556	227557
Pre-reform treated mean	24.807	.116			25.02	.12		
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Criminal history FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Includes all felony treated and all untreated probation spells beginning 1-3 years before the reform and 0-2 years afterwards. Column 1 reports the difference-in-differences effect of the reform on total days incarcerated for technical violations that occur in the first year of a spell for black offenders. Column 2 reports the effect on any incarceration for technical violations. Columns 3 and 4 report the 2SLS effects, i.e., the reduction in incarceration days and any incarceration for offenders no longer revoked in their first year due to the reform. Columns 5-8 repeat the same regressions for non-black offenders. Post is indicator for starting probation after December 1, 2011, the date JRA reforms took effect. Demographic controls include five-year age bins and indicators for race and gender. Criminal history controls include fixed effects criminal history points and prior sentences to supervised probation or incarceration. Standard errors are clustered by individual.

77

Table A13: Effect of Reform on Misdemeanants Unaffected by CRVs

	Black		Non-black	
	(1)	(2)	(3)	(4)
	Revoke	Arrest	Revoke	Arrest
Post-reform	-0.00432***	-0.0202***	-0.00292***	-0.00147
	(0.000848)	(0.00445)	(0.000489)	(0.00292)
Treated	0.174***	-0.0461***	0.121***	-0.00664*
	(0.00283)	(0.00412)	(0.00210)	(0.00319)
Post-x-treat	-0.0795***	0.0259***	-0.0282***	0.0178***
	(0.00399)	(0.00664)	(0.00309)	(0.00494)
N	78124	78124	128281	128281
Pre-reform treated mean	.191	.299	.136	.254
Demographic controls	Yes	Yes	Yes	Yes
Criminal history FE	Yes	Yes	Yes	Yes
Accuracy		.388 (.083)		.899 (.184)
False negative		.908 (.019)		.913 (.016)
False positive		.086 (.011)		.004 (.008)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Includes all misdemeanor treated and all untreated probation spells beginning 1-2 years before the reform or in 2016, after CRVs were eliminated for misdemeanor probationers by the legislature. Post is indicator for starting probation after December 1, 2011, the date JRA reforms took effect. Demographic controls include five-year age bins and indicators for race and gender. Criminal history controls include fixed effects criminal history points and prior sentences to supervised probation or incarceration. Standard errors are clustered by individual.

Table A14: Effect of Reform With Longer Reoffending Window

	Black		Non-black	
	(1)	(2)	(3)	(4)
	Revoke	Arrest	Revoke	Arrest
Post-reform	-0.00380***	-0.0521***	-0.00139***	-0.0261***
	(0.000308)	(0.00176)	(0.000190)	(0.00118)
Treated	0.160***	-0.0562***	0.110***	0.00143
	(0.000985)	(0.00163)	(0.000739)	(0.00125)
Post-x-treat	-0.0840***	0.0433***	-0.0339***	0.0208***
	(0.00123)	(0.00232)	(0.000995)	(0.00176)
N	625948	625948	977765	977765
Pre-reform treated mean	.181	.344	.131	.293
Demographic controls	Yes	Yes	Yes	Yes
Criminal history FE	Yes	Yes	Yes	Yes
Accuracy		.54 (.027)		.802 (.054)
False negative		.886 (.005)		.921 (.005)
False positive		.076 (.004)		.011 (.003)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Includes all treated and all untreated probation spells beginning 1-3 years before the reform and 0-2 years afterwards. Reoffending is measured of 455 days (instead of 365) to allow for potential incapacitation effects of CRVs. Post is indicator for starting probation after December 1, 2011, the date JRA reforms took effect. Demographic controls include five-year age bins and indicators for race and gender. Criminal history controls include fixed effects criminal history points and prior sentences to supervised probation or incarceration. Standard errors are clustered by individual.

Table A15: Effect of Reform by Crime Type

	Black			Non-black		
	(1)	(2)	(3)	(4)	(5)	(6)
	Any	Misd/fel	Fel	Any	Misd/fel	Fel
Post-reform	-0.0111***	-0.00925***	0.00208	-0.00661***	-0.00188	0.00325***
	(0.00281)	(0.00274)	(0.00168)	(0.00190)	(0.00178)	(0.000963)
Treated	-0.0467***	-0.0411***	-0.00294	-0.000306	0.00161	0.00738***
	(0.00268)	(0.00262)	(0.00163)	(0.00207)	(0.00195)	(0.00110)
Post-x-treat	0.0237***	0.0211***	0.00578*	0.0182***	0.0181***	0.00936***
	(0.00383)	(0.00374)	(0.00237)	(0.00295)	(0.00279)	(0.00163)
N	217222	217222	217222	328784	328784	328784
Pre-reform treated mean	.314	.29	.092	.264	.226	.062
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Criminal history FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: 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 December 1, 2011, the date JRA reforms took effect. Demographic controls include five-year age bins and indicators for race and gender. Criminal history controls include fixed effects criminal history points and prior sentences to supervised probation or incarceration. Standard errors are clustered by individual.

Table A16: Impact of Data Window on Effects of Reform

	Non-black			Black		
	Revoked					
	(1)	(2)	(3)	(4)	(5)	(6)
	1yr	2yr	3yr	1yr	2yr	3yr
Post-reform	-0.0013**	-0.00087**	-0.00064*	-0.0048***	-0.0041***	-0.0040***
	(0.00048)	(0.00034)	(0.00028)	(0.00077)	(0.00054)	(0.00044)
Treated	0.12***	0.11***	0.11***	0.16***	0.16***	0.16***
	(0.0018)	(0.0013)	(0.0010)	(0.0024)	(0.0017)	(0.0014)
Post-x-treat	-0.042***	-0.037***	-0.036***	-0.070***	-0.076***	-0.079***
	(0.0025)	(0.0017)	(0.0014)	(0.0031)	(0.0022)	(0.0018)
R-squared	0.081	0.079	0.078	0.090	0.091	0.092
Pre-reform treated mean	.136	.131	.128	.181	.181	.182
	Arrest					
Post-reform	-0.0036	-0.0066***	-0.0081***	-0.0036	-0.011***	-0.019***
	(0.0026)	(0.0019)	(0.0016)	(0.0039)	(0.0028)	(0.0024)
Treated	-0.0041	-0.00031	0.0019	-0.044***	-0.047***	-0.049***
	(0.0029)	(0.0021)	(0.0017)	(0.0038)	(0.0027)	(0.0022)
Post-x-treat	0.021***	0.018***	0.018***	0.016**	0.024***	0.029***
	(0.0041)	(0.0029)	(0.0024)	(0.0054)	(0.0038)	(0.0032)
N	165936	328784	488779	109764	217222	319596
R-squared	0.072	0.073	0.072	0.083	0.080	0.079
Pre-reform treated mean	.257	.264	.268	.31	.314	.317
Accuracy	.517	.543	.58	.205	.306	.363
False negative rate	.923	.93	.929	.956	.931	.917
False positive rate	.032	.027	.024	.099	.094	.091
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Criminal history FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Includes all treated and untreated probation spells beginning within 1, 2, and 3 years before the reform and within 0, 1, and 2 afterwards, as indicated in the column header. Post is indicator for starting probation after December 1, 2011, the date JRA reforms took effect. Demographic controls include five-year age bins and indicators for race and gender. Criminal history controls include fixed effects criminal history points and prior sentences to supervised probation or incarceration. Standard errors are clustered by individual.

Table A17: Dynamic Decomposition of Racial Gaps in Technical Revocations

	Overall rates		Decomposition	
	Non-black	Black	Difference	Share of gap
Probability of technical revoke in 1080 days				
$Pr(R_i(0) \leq 1080)$	0.045	0.100	0.056	100.0%
Distribution of risk				
$Pr(Y_i(0) \leq 360)$	0.313	0.363	0.05	6.5%
$Pr(Y_i(0) \leq 720)$	0.426	0.488	0.061	10.1%
$Pr(Y_i(0) \leq 1080)$	0.498	0.558	0.060	11.0%
$Pr(Y_i(0) > 1080)$	0.502	0.442	-0.060	-9.6%
Total contribution				1.5%
Probability of revoke conditional on risk				
$Pr(R_i(0) < 360\|Y_i(0) < 360)$	0.070	0.077	0.007	4.5%
$Pr(R_i(0) < 720\|Y_i(0) < 720)$	0.063	0.106	0.043	34.6%
$Pr(R_i(0) < 1080\|Y_i(0) < 1080)$	0.073	0.110	0.037	34.3%
$Pr(R_i(0) < 1080\|Y_i(0) \geq 1080)$	0.017	0.088	0.072	64.3%
Total contribution				98.5%

Notes: Table decomposes the difference in the risk of revocation for technical violations between black and white probationers into the contributions of differences in arrest risk and differences in the likelihood of revocation conditional on arrest risk using the multi-period model described in Section 3. The decomposition applies to the population with $R_i(1) > 1080$, or "potential compliers." These are individuals who are not revoked for breaking rules within three yeras even after the reform. The first row reports the share of white and black offenders caught by the drug and administrative rules affected by the reform and the black rate minus the white rate. The remainder of the table decomposes this differences into the share explained by differences in $Pr(Y_i(0) = k)$ and differences in revocation conditional on $Y_i(0)$. The rows under "Distribution of Risk" show the share of potential compliers by race with $Y_i(0)$ in certain ranges, the black-white gap, and the contribution of this gap to the total disparity. The rows under "Probability of revoke conditional on risk" show mean values of technical revocation rates for potential compliers with $Y_i(0)$ in certain ranges, the gap, and the contribution of this gap to the total disparity. Since crime is measured up to a max of a 3 year horizon, risk distributions are not observed beyond this point. Y_i is therefore binned in 90-day intervals up to 3 years with a final bin reflecting 3 years or later. Additional details are available in Section A4.

Table A18: Rule Violations By Probation Outcome Post Reform

	Reporting	Drug	Fees	Other
Non-black probationers				
Arrest	0.10 (0.00)	0.10 (0.00)	0.09 (0.00)	0.04 (0.00)
Incar for TVs	0.07 (0.00)	0.07 (0.00)	0.03 (0.00)	0.03 (0.00)
Successful completion	0.09 (0.00)	0.11 (0.00)	0.20 (0.00)	0.04 (0.00)
Black probationers				
Arrest	0.11 (0.00)	0.12 (0.00)	0.13 (0.00)	0.04 (0.00)
Incar for TVs	0.06 (0.00)	0.09 (0.00)	0.05 (0.00)	0.04 (0.00)
Successful completion	0.14 (0.00)	0.17 (0.00)	0.36 (0.00)	0.05 (0.00)

Notes: Table reports shares of probationers ever breaking rules of given types *prior* to finishing their spell broken down by reason for spell exit and race. Probationers can exit due to an arrest, incarceration for rule violations, or sucessfuly completing their supervision spell. For example, the first row reports the share of white probationers exiting probation due to a criminal arrest who break reporting, drug, fees, and other rules prior to their exit. If not for censoring due to incarceration for rule violations, these shares would reflect the true and false positive rates associated with using each rule type as signals of arrest risk. Rule violations are broken 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. Volations are coded 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.

Table A19: Mixture Model Parameter Estimates for Men

	Black men		Non-black men	
	Arrest	Revoke	Arrest	Revoke
Duration	-0.17 (0.11)	3.78 (0.17)	-0.87 (0.10)	2.86 (0.20)
Duration^2	-1.85 (0.73)	-21.78 (1.26)	2.16 (0.68)	-18.56 (1.54)
Duration^3	4.98 (1.86)	42.21 (3.53)	-4.40 (1.75)	36.76 (4.40)
Duration^4	-4.79 (2.03)	-37.99 (4.16)	4.87 (1.93)	-33.68 (5.21)
Duration^5	1.56 (0.79)	12.94 (1.72)	-2.08 (0.76)	11.63 (2.17)
Has 2 spells	0.85 (0.01)	0.77 (0.02)	1.21 (0.01)	1.09 (0.02)
Second spell	-0.19 (0.03)	0.09 (0.04)	-0.34 (0.03)	-0.03 (0.05)
Second spell x dur.	-0.13 (0.12)	-0.02 (0.22)	-0.02 (0.12)	0.22 (0.21)
Second spell x dur.^{2}	0.60 (0.71)	-1.56 (1.36)	-0.13 (0.65)	-2.86 (1.27)
Second spell x dur.^{3}	-1.42 (1.73)	4.96 (3.60)	0.14 (1.57)	8.22 (3.30)
Second spell x dur.^{4}	1.38 (1.85)	-5.57 (4.15)	0.08 (1.67)	-9.01 (3.75)
Second spell x dur.^{5}	-0.46 (0.72)	2.11 (1.71)	-0.11 (0.64)	3.43 (1.53)
Calendar time	-0.02 (0.01)	-0.22 (0.02)	0.05 (0.01)	-0.05 (0.02)
Calendar time^2	-0.00 (0.01)	-0.15 (0.01)	0.02 (0.01)	-0.08 (0.01)
Age	-2.50 (0.13)	-3.35 (0.20)	-2.91 (0.13)	-2.07 (0.22)
Age^2	4.14 (0.28)	6.67 (0.43)	5.50 (0.27)	4.40 (0.48)
Age^3	-2.03 (0.16)	-3.49 (0.24)	-2.90 (0.15)	-2.53 (0.26)
Post reform	0.05 (0.01)	-0.51 (0.03)	0.04 (0.01)	-0.40 (0.03)
Type locations				
Type 1	-6.92 (0.00)	-7.02 (0.08)	-7.72 (0.00)	-8.55 (0.20)
Type 2	-5.43 (0.00)	-7.25 (0.09)	-5.87 (0.00)	-8.17 (0.16)
Type 3	-5.41 (0.00)	-5.46 (0.08)	-5.82 (0.00)	-6.27 (0.09)
Type 4	-3.45 (0.06)	-5.98 (0.19)	-3.72 (0.05)	-6.61 (0.24)
Type shares				
Type 1	0.12 (0.01)		0.06 (0.00)	
Type 2	0.58 (0.03)		0.58 (0.04)	
Type 3	0.23 (0.03)		0.30 (0.04)	
Type 4	0.08 (0.00)		0.06 (0.00)	
Total spells	173,441		207,388	
Total individuals	139,373		174,775	
Log likelihood	-715877.466		-739260.018	

Notes: Table reports estimates of the mixed logit model described in Section 5. Duration, age, and calendar time are standardized (s.d. 1 and mean 0) in estimation. Standard errors are the robust "sandwich form" clustered by individual. Hazards are discreteized into 7-day units. Given the logit formulation for the hazard, coefficients can therefore be interpreted as effects on the weekly hazard log odds.

Table A20: Mixture Model Parameter Estimates for Women

	Black women		Non-black women	
	Arrest	Incar for TVs	Arrest	Incar for TVs
Duration	-0.51 (1.68)	3.74 (1.95)	-0.74 (0.18)	2.72 (0.32)
Duration^2	1.31 (6.82)	-21.52 (6.58)	0.96 (1.15)	-21.03 (2.29)
Duration^3	-2.60 (12.19)	40.43 (11.96)	-1.16 (2.92)	47.74 (6.50)
Duration^4	2.63 (10.46)	-35.15 (11.31)	1.29 (3.19)	-48.69 (7.81)
Duration^5	-1.05 (3.47)	11.57 (4.25)	-0.66 (1.25)	18.21 (3.31)
Has 2 spells	1.25 (0.03)	1.08 (0.07)	1.33 (0.02)	1.29 (0.04)
Second spell	-0.30 (0.09)	0.02 (0.19)	-0.38 (0.05)	0.02 (0.07)
Second spell x dur.	-0.04 (0.33)	-0.08 (0.62)	-0.25 (0.19)	-0.27 (0.33)
Second spell x dur.^{2}	-1.08 (2.32)	-2.14 (3.09)	0.89 (1.09)	-0.12 (2.08)
Second spell x dur.^{3}	3.48 (5.40)	8.73 (7.66)	-1.55 (2.58)	1.25 (5.55)
Second spell x dur.^{4}	-4.03 (5.32)	-12.41 (8.60)	1.17 (2.71)	-1.46 (6.40)
Second spell x dur.^{5}	1.62 (1.91)	5.83 (3.48)	-0.31 (1.04)	0.54 (2.63)
Calendar time	0.01 (0.02)	-0.14 (0.05)	0.07 (0.01)	0.07 (0.03)
Calendar time^2	0.00 (0.02)	-0.08 (0.03)	0.01 (0.01)	0.03 (0.02)
Age	-1.79 (0.27)	-3.91 (0.53)	-0.37 (0.22)	-0.08 (0.43)
Age^2	3.24 (0.57)	8.36 (1.04)	0.97 (0.46)	0.84 (0.91)
Age^3	-1.71 (0.32)	-4.56 (0.56)	-0.84 (0.25)	-1.01 (0.50)
Post reform	0.04 (0.03)	-0.56 (0.06)	0.05 (0.02)	-0.40 (0.05)
Type locations				
Type 1	-7.99 (0.53)	-8.43 (1.05)	-8.14 (0.00)	-8.74 (0.46)
Type 2	-6.18 (0.14)	-5.39 (3.21)	-5.97 (0.00)	-7.92 (0.07)
Type 3	-5.89 (0.01)	-7.67 (1.60)	-5.80 (0.01)	-5.69 (0.16)
Type 4	-3.59 (1.80)	-6.38 (3.03)	-3.60 (0.06)	-6.63 (0.73)
Type shares				
Type 1	0.12 (0.03)		0.06 (0.01)	
Type 2	0.09 (0.36)		0.79 (0.02)	
Type 3	0.73 (0.44)		0.09 (0.02)	
Type 4	0.06 (0.05)		0.05 (0.00)	
Total spells	53,258		78,695	
Total individuals	45,670		67,003	
Log likelihood	-181267.502		-265467.568	

Notes: Table reports estimates of the mixed logit model described in Section 5. Duration, age, and calendar time are standardized (s.d. 1) in estimation. Standard errors are the robust "sandwich form" clustered by individual. Hazards are discreteized into 30 day units. Given the logit formulation for the hazard, coefficients can therefore be interpreted as effects on the monthly hazard log odds.

Table A21: Continuous Heterogeneity Model Parameter Estimates for Men

	Black men		Non-black men	
	Arrest	Incar for TVs	Arrest	Incar for TVs
Duration	-0.71 (0.09)	3.93 (0.16)	-1.43 (0.09)	3.07 (0.18)
Duration^2	0.43 (0.66)	-22.36 (1.22)	4.34 (0.66)	-19.27 (1.35)
Duration^3	0.75 (1.74)	43.30 (3.41)	-8.37 (1.73)	37.89 (3.82)
Duration^4	-1.12 (1.93)	-38.94 (4.02)	8.33 (1.91)	-34.57 (4.53)
Duration^5	0.34 (0.77)	13.25 (1.67)	-3.22 (0.75)	11.90 (1.89)
Has 2 spells	0.82 (0.01)	0.78 (0.02)	1.16 (0.01)	1.11 (0.02)
Second spell	-0.20 (0.03)	0.09 (0.04)	-0.34 (0.03)	-0.04 (0.05)
Second spell x dur.	-0.16 (0.12)	0.02 (0.21)	-0.06 (0.12)	0.29 (0.20)
Second spell x dur.^{2}	0.83 (0.71)	-1.79 (1.32)	0.14 (0.66)	-3.12 (1.25)
Second spell x dur.^{3}	-1.85 (1.73)	5.47 (3.50)	-0.34 (1.58)	8.71 (3.26)
Second spell x dur.^{4}	1.74 (1.86)	-6.06 (4.04)	0.47 (1.68)	-9.44 (3.70)
Second spell x dur.^{5}	-0.58 (0.72)	2.28 (1.66)	-0.23 (0.65)	3.57 (1.51)
Calendar time	-0.02 (0.01)	-0.23 (0.02)	0.05 (0.01)	-0.05 (0.02)
Calendar time^2	0.00 (0.01)	-0.15 (0.01)	0.02 (0.01)	-0.09 (0.01)
Age	-2.53 (0.13)	-3.36 (0.20)	-2.83 (0.12)	-2.12 (0.23)
Age^2	4.21 (0.28)	6.70 (0.44)	5.34 (0.26)	4.50 (0.49)
Age^3	-2.06 (0.16)	-3.51 (0.24)	-2.81 (0.14)	-2.58 (0.27)
Post reform	0.05 (0.01)	-0.51 (0.03)	0.03 (0.01)	-0.40 (0.03)
σ, ρ				
Arrest	0.66 (0.01)	0.20 (0.03)	0.54 (0.01)	0.33 (0.03)
Tech. Incar.		0.96 (0.02)		1.06 (0.03)
Total spells	173,441		207,388	
Total individuals	139,373		174,775	
Log likelihood	-716000.129		-739434.749	

Notes: Table reports estimates of the mixed logit model described in Section 5. Duration, age, and calendar time are standardized (s.d. 1) in estimation. Standard errors are the robust "sandwich form" clustered by individual. Hazards are discreteized into 7 day units. Given the logit formulation for the hazard, coefficients can therefore be interpreted as effects on the weekly hazard log odds. Unobserved heterogeneity across the two risks is bivariate normal. The σ, ρ estimates correspond to the square root of the variance and correlations of each component.

Table A22: Continuous Heterogeneity Model Parameter Estimates for Women

	Black women		Non-black women	
	Arrest	Incar for TVs	Arrest	Incar for TVs
Duration	-1.14 (0.19)	3.91 (0.37)	-1.39 (0.16)	2.68 (0.31)
Duration^2	3.61 (1.36)	-22.37 (2.68)	3.59 (1.10)	-21.05 (2.34)
Duration^3	-6.60 (3.53)	42.20 (7.58)	-6.02 (2.88)	48.04 (6.77)
Duration^4	5.99 (3.87)	-36.81 (9.03)	5.53 (3.17)	-49.14 (8.17)
Duration^5	-2.14 (1.52)	12.15 (3.80)	-2.07 (1.25)	18.40 (3.47)
Has 2 spells	1.20 (0.02)	1.08 (0.04)	1.27 (0.02)	1.27 (0.04)
Second spell	-0.31 (0.06)	0.02 (0.10)	-0.39 (0.04)	0.03 (0.07)
Second spell x dur.	-0.11 (0.26)	-0.02 (0.46)	-0.28 (0.19)	-0.28 (0.33)
Second spell x dur.^{2}	-0.61 (1.43)	-2.50 (2.86)	1.21 (1.07)	-0.16 (2.09)
Second spell x dur.^{3}	2.60 (3.40)	9.54 (7.58)	-2.21 (2.55)	1.45 (5.58)
Second spell x dur.^{4}	-3.29 (3.57)	-13.23 (8.70)	1.77 (2.69)	-1.72 (6.42)
Second spell x dur.^{5}	1.38 (1.36)	6.13 (3.54)	-0.51 (1.03)	0.65 (2.64)
Calendar time	0.01 (0.02)	-0.14 (0.05)	0.06 (0.02)	0.07 (0.03)
Calendar time^2	0.01 (0.01)	-0.08 (0.03)	0.01 (0.01)	0.03 (0.02)
Age	-1.82 (0.26)	-3.92 (0.47)	-0.28 (0.21)	-0.07 (0.43)
Age^2	3.31 (0.55)	8.40 (1.00)	0.77 (0.44)	0.82 (0.91)
Age^3	-1.74 (0.30)	-4.59 (0.54)	-0.72 (0.24)	-1.00 (0.50)
Post reform	0.03 (0.03)	-0.57 (0.06)	0.05 (0.02)	-0.40 (0.05)
σ, ρ				
Arrest	0.72 (0.02)	0.22 (0.06)	0.53 (0.02)	0.34 (0.06)
Tech. Incar.		1.24 (0.09)		1.09 (0.08)
Total spells	53,258		78,695	
Total individuals	45,670		67,003	
Log likelihood	-181323.295		-265536.900	

Notes: Table reports estimates of the mixed logit model described in Section 5. Duration, age, and calendar time are standardized (s.d. 1) in estimation. Standard errors are the robust "sandwich form" clustered by individual. Hazards are discreteized into 7 day units. Given the logit formulation for the hazard, coefficients can therefore be interpreted as effects on the weekly hazard log odds. Unobserved heterogeneity across the two risks is bivariate normal. The σ, ρ estimates correspond to the square root of the variance and correlations of each component.

Table A23: Mixture Model With Multiple Violation Types Parameter Estimates for Black Men

| | Black men | | | | | |
	Arrest	Reporting	Drug	Fees/Fines	Other	Revoke | viol
Duration	-0.41 (0.38)	3.85 (0.22)	6.78 (0.21)	9.76 (0.28)	-0.44 (0.31)	-1.35 (0.22)
Duration^2	-0.64 (1.78)	-20.65 (1.24)	-31.41 (1.34)	-35.44 (1.51)	-1.23 (2.30)	-0.21 (1.57)
Duration^3	2.46 (3.68)	39.03 (3.14)	58.43 (3.45)	55.02 (3.45)	5.88 (6.17)	4.58 (4.23)
Duration^4	-2.46 (3.52)	-34.64 (3.53)	-51.59 (3.83)	-41.12 (3.50)	-8.99 (7.00)	-5.54 (4.90)
Duration^5	0.75 (1.26)	11.71 (1.43)	17.42 (1.52)	12.01 (1.31)	4.29 (2.82)	1.96 (2.03)
Has 2 spells	0.83 (0.03)	0.58 (0.02)	0.49 (0.03)	0.28 (0.02)	0.50 (0.04)	
Second spell	-0.20 (0.06)	0.15 (0.04)	-0.15 (0.07)	-0.05 (0.12)	0.12 (0.07)	
Second spell x dur.	-0.11 (0.26)	-0.01 (0.19)	0.22 (0.26)	-0.42 (0.38)	-0.38 (0.37)	
Second spell x dur.^{2}	0.66 (1.27)	-1.76 (1.10)	-1.42 (1.37)	1.52 (1.67)	2.17 (2.19)	
Second spell x dur.^{3}	-1.64 (2.83)	5.84 (2.81)	3.53 (3.28)	-2.22 (3.50)	-4.89 (5.50)	
Second spell x dur.^{4}	1.62 (2.86)	-6.70 (3.15)	-3.50 (3.50)	1.42 (3.41)	4.82 (6.00)	
Second spell x dur.^{5}	-0.55 (1.06)	2.61 (1.27)	1.19 (1.36)	-0.33 (1.24)	-1.71 (2.35)	
Calendar time	-0.04 (0.03)	-0.06 (0.02)	0.06 (0.02)	0.07 (0.01)	0.31 (0.03)	
Calendar time^2	-0.01 (0.01)	0.04 (0.01)	-0.11 (0.01)	-0.14 (0.01)	0.07 (0.02)	
Age	-2.45 (0.13)	-2.52 (0.18)	-1.15 (0.22)	-0.81 (0.19)	-5.00 (0.39)	-1.44 (0.23)
Age^2	4.08 (0.28)	5.19 (0.38)	1.60 (0.47)	2.11 (0.39)	9.31 (0.85)	2.79 (0.50)
Age^3	-2.01 (0.16)	-2.87 (0.21)	-0.64 (0.26)	-1.30 (0.21)	-4.65 (0.47)	-1.39 (0.27)
Post reform	0.05 (0.06)	-0.08 (0.02)	0.00 (0.03)	-0.01 (0.02)	-0.29 (0.05)	
Num. prev. viol.						0.04 (0.01)
Constant						-0.27 (0.03)
Drug viol.						-0.72 (0.02)
Fees viol.						-1.27 (0.02)
Other viol.						-1.23 (0.03)
Post x rep. viol.						-0.64 (0.02)
Post x drug viol.						-1.38 (0.03)
Post x fees viol.						-1.42 (0.04)
Post x other viol.						-1.53 (0.07)
Type locations						
Type 1	-5.82 (0.02)	-7.14 (0.10)	-7.40 (0.06)	-6.09 (0.03)	-8.08 (0.05)	
Type 2	-5.42 (0.02)	-5.59 (0.04)	-6.66 (0.25)	-7.02 (0.10)	-8.60 (0.22)	
Type 3	-5.31 (0.05)	-6.32 (0.06)	-5.17 (0.08)	-6.51 (0.24)	-6.88 (0.48)	
Type 4	-4.00 (0.26)	-5.18 (0.24)	-5.88 (0.11)	-5.74 (0.10)	-5.88 (0.08)	
Type shares						
Type 1	0.46 (0.04)					
Type 2	0.27 (0.02)					
Type 3	0.14 (0.02)					
Type 4	0.13 (0.02)					
Total spells	173,441					
Total individuals	139,373					
Log likelihood	-1316925.688					

Notes: Table reports estimates of the mixed logit model described in Section 5 when decomposing incarceration risk across violation types. Duration, age, and calendar time are standardized (s.d. 1 and mean 0) in estimation. Standard errors are the robust "sandwich form" clustered by individual. Hazards are discreteized into 7-day units. Given the logit formulation for the hazard, coefficients can therefore be interpreted as effects on the weekly hazard log odds.

Table A24: Mixture Model With Multiple Violation Types Parameter Estimates for White Men

| | Non-black men | | | | | |
	Arrest	Reporting	Drug	Fees/Fines	Other	Revoke | viol
Duration	-1.07 (0.10)	3.23 (0.15)	5.95 (0.21)	9.82 (0.31)	-1.34 (0.29)	-1.43 (0.23)
Duration^2	3.22 (0.69)	-19.35 (1.10)	-26.19 (1.43)	-37.45 (1.76)	6.79 (2.16)	0.49 (1.68)
Duration^3	-6.62 (1.78)	37.69 (3.02)	46.23 (3.74)	60.76 (4.09)	-17.16 (5.92)	3.58 (4.61)
Duration^4	6.96 (1.96)	-34.11 (3.50)	-39.64 (4.18)	-47.52 (4.22)	16.50 (6.81)	-4.96 (5.40)
Duration^5	-2.80 (0.77)	11.70 (1.44)	13.23 (1.67)	14.51 (1.59)	-5.39 (2.78)	1.81 (2.26)
Has 2 spells	1.22 (0.01)	0.85 (0.02)	1.00 (0.02)	0.59 (0.02)	0.77 (0.03)	
Second spell	-0.35 (0.03)	0.12 (0.04)	-0.26 (0.07)	-0.23 (0.14)	0.05 (0.08)	
Second spell x dur.	0.01 (0.12)	-0.14 (0.18)	0.27 (0.25)	-0.22 (0.39)	-0.08 (0.35)	
Second spell x dur.^{2}	-0.13 (0.67)	-0.50 (1.04)	-1.73 (1.34)	1.06 (1.71)	-0.30 (2.09)	
Second spell x dur.^{3}	0.08 (1.61)	2.67 (2.65)	3.92 (3.19)	-1.59 (3.61)	1.10 (5.25)	
Second spell x dur.^{4}	0.16 (1.71)	-3.51 (2.95)	-3.27 (3.39)	1.08 (3.53)	-0.87 (5.75)	
Second spell x dur.^{5}	-0.14 (0.66)	1.47 (1.18)	0.88 (1.31)	-0.29 (1.29)	0.16 (2.27)	
Calendar time	0.04 (0.01)	0.02 (0.02)	0.14 (0.02)	0.15 (0.02)	0.29 (0.03)	
Calendar time^2	0.01 (0.01)	0.07 (0.01)	-0.07 (0.01)	-0.14 (0.01)	0.08 (0.02)	
Age	-2.88 (0.13)	-0.72 (0.19)	-2.94 (0.23)	-1.44 (0.22)	-4.34 (0.37)	-0.65 (0.26)
Age^2	5.43 (0.27)	1.82 (0.41)	5.42 (0.50)	3.15 (0.46)	8.25 (0.78)	1.40 (0.55)
Age^3	-2.86 (0.15)	-1.26 (0.23)	-2.75 (0.27)	-1.74 (0.25)	-4.12 (0.43)	-0.86 (0.30)
Post reform	0.04 (0.02)	0.05 (0.02)	-0.05 (0.03)	-0.05 (0.02)	-0.21 (0.05)	
Num. prev. viol.						-0.00 (0.02)
Constant						-0.35 (0.03)
Drug viol.						-0.70 (0.02)
Fees viol.						-1.19 (0.03)
Other viol.						-1.24 (0.03)
Post x rep. viol.						-0.40 (0.02)
Post x drug viol.						-1.21 (0.04)
Post x fees viol.						-1.25 (0.05)
Post x other viol.						-1.45 (0.07)
Type locations						
Type 1	-6.25 (0.01)	-6.33 (0.07)	-6.59 (0.10)	-8.37 (0.17)	-8.84 (0.23)	
Type 2	-6.25 (0.01)	-8.37 (0.09)	-8.52 (0.08)	-7.48 (0.05)	-9.08 (0.07)	
Type 3	-5.56 (0.01)	-6.64 (0.06)	-7.04 (0.10)	-6.22 (0.04)	-7.52 (0.07)	
Type 4	-4.32 (0.03)	-5.69 (0.05)	-5.14 (0.04)	-6.91 (0.12)	-6.42 (0.06)	
Type shares						
Type 1	0.17 (0.01)					
Type 2	0.40 (0.02)					
Type 3	0.32 (0.01)					
Type 4	0.11 (0.00)					
Total spells	207,388					
Total individuals	174,775					
Log likelihood	-1285767.598					

Notes: Table reports estimates of the mixed logit model described in Section 5 when decomposing incarceration risk across violation types. Duration, age, and calendar time are standardized (s.d. 1 and mean 0) in estimation. Standard errors are the robust "sandwich form" clustered by individual. Hazards are discreteized into 7-day units. Given the logit formulation for the hazard, coefficients can therefore be interpreted as effects on the weekly hazard log odds.

Table A25: Estimates of lower and upper bounds of the costs/value of crime

| Offense category | Lower bound $ | | | Upper bound $ | | |
	Raw estimate	Including discounting	Reference	Raw estimate	Including discounting	Reference
Homicide	7,000,000	7,350,000	Chalfin and McCrary (2017)	9,700,000	19,205,337	Cohen et al. (2011)
Rape	142,020	149,121	Chalfin and McCrary (2017)	237,000	469,243.8	Cohen et al. (2011)
Assault	38,924	40,870.2	Chalfin and McCrary (2017)	70,000	138,595.2	Cohen et al. (2011)
Robbery	12,624	13,255.2	Chalfin and McCrary (2017)	232,000	459,344.1	Cohen et al. (2011)
Arson	38,000	128,681	Miller, Cohen and Wiersema (1996)	38,000	128,681	Miller, Cohen and Wiersema (1996)
Burglary	2,104	2,209.2	Chalfin and McCrary (2017)	25,000	49498.29	Cohen et al. (2011)
Larceny	473	497	Chalfin and McCrary (2017)	370	1,253	Miller, Cohen and Wiersema (1996)
Theft	473	497	Chalfin and McCrary (2017)	370	1,253	Miller, Cohen and Wiersema (1996)
Drug	500	990		2,544	2,945	Mueller-Smith (2015)
DWI	500	990		25,842	29,915	Mueller-Smith (2015)
Other	500	990	Cohen et al. (2011)	500	990	Cohen et al. (2011)

Notes: "Discounting" means updating the cost estimate to 2018 $, using a rate of 5% as in Mueller-Smith (2015). Offenses without a relevant cost estimate are assigned a value of $990 (in 2018 $) as was suggested by Cohen et al. (2011). The lower bounds for drug and DWI offenses were assigned in this way. Note that only Miller, Cohen and Wiersema (1996) and Cohen et al. (2011) calculated value of crime estimates the other studies used estimates from various other studies including from Miller, Cohen and Wiersema (1996) and Cohen et al. (2011).