THE ROADBLOCK TO REFORM

BY MEGAN T. STEVENSON AND JENNIFER L. DOLEAC
Algorithmic risk assessment tools have become a popular element of criminal justice reforms, often with the explicit goal of reducing incarceration rates. The hope is that these data-driven tools will standardize decisions about pretrial detention and sentencing, and ensure that only the most high-risk offenders are incarcerated. However, the effects of these tools depend crucially on how judges use them.

This brief considers judicial reforms in Kentucky and Virginia as case studies of the effects of algorithmic risk assessment tools in practice. We show that, in both states, reforms aimed at reducing incarceration for low-risk offenders had little to no impact on incarceration rates. While these tools clearly recommended less incarceration for a large share of defendants, they had little effect on judges’ incarceration decisions. However, there is tremendous variation across judges in how closely they follow the risk assessment recommendations. We discuss what these results mean in terms of how to affect meaningful criminal justice reform, including ways citizens and policymakers can align judicial incentives so that reforms have their intended effects.
On August 28, 2018, California’s governor signed a major criminal justice reform bill into law. This bill will replace monetary bail with a system in which pretrial custody is determined partially by an algorithmic evaluation of the likelihood that a defendant will commit future crime or fail to appear in court. Instead, the pretrial release rate for low and moderate risk defendants increased by only four percentage points, from 63% to 67%.

The median judge in Kentucky grants release without monetary bail to only 37% of defendants with low or moderate risk status. In other words, the median judge overrules the presumptive default associated with the risk assessment about 2/3 of the time.

If judges had followed the recommendations associated with the risk assessment, there would have been up to a 25% increase in the diversion rate among risk-assessment-eligible offenders after risk assessment was adopted statewide. Instead, the diversion rate remained virtually unchanged, dipping just slightly from 34% to 33%.

The median judge in Virginia diverts only about 40% of those who were recommended for diversion by the risk assessment instrument.

**KEY FINDINGS:**

- If Kentucky judges had followed the recommendations associated with the risk assessment in all cases, the pretrial release rate among low and moderate risk defendants would have jumped up by 37 percentage points after risk assessment was made mandatory. Instead, the pretrial release rate for low and moderate risk defendants increased by only four percentage points, from 63% to 67%.

- The median judge in Kentucky grants release without monetary bail to only 37% of defendants with low or moderate risk status. In other words, the median judge overrules the presumptive default associated with the risk assessment about 2/3 of the time.

- If judges had followed the recommendations associated with the risk assessment, there would have been up to a 25% increase in the diversion rate among risk-assessment-eligible offenders after risk assessment was adopted statewide. Instead, the diversion rate remained virtually unchanged, dipping just slightly from 34% to 33%.

- The median judge in Virginia diverts only about 40% of those who were recommended for diversion by the risk assessment instrument.

**BACKGROUND:**

**Algorithmic Risk Assessment as Criminal Justice Reform**

On August 28, 2018, California’s governor signed a major criminal justice reform bill into law. This bill will replace monetary bail with a system in which pretrial custody is determined partially by an algorithmic evaluation of the likelihood that a defendant will commit future crime or fail to appear in court. California is following in the footsteps of many other jurisdictions that have implemented criminal justice reform by way of algorithmic risk evaluations. In recent years, pretrial risk assessment has been adopted in dozens of jurisdictions and at least six entire states. Risk assessment is used in sentencing in at least twenty-eight full states; at least seven additional states have at least one county that uses risk assessment tools at sentencing. It is also widely used in determining whether a prisoner will be released on parole, in determining supervision levels for probation, and in selecting prison security level.

Risk assessment algorithms are designed to predict the likelihood of committing new crime or failing to appear in court. Built by data scientists, risk assessments are designed by estimating the correlation between “inputs” —criminal history, education, employment, age, gender, etc. — and the outcome that the tool is trying to predict. These statistical correlations will then be converted into an algorithm, in which points are added or subtracted from the total risk score based on the input values. A person will be classified as “high risk” if their total points exceed a certain cutoff.

Proponents of risk assessment have argued that they will help to lower incarceration rates with minimal effect on public safety. In fact, jurisdictions often adopt risk assessment with the stated goal of lowering jail and prison populations. The risk evaluations are paired with explicit action recommendations: for example, a jurisdiction might recommend release or diversion for all defendants who are ranked low or moderate risk. However, the final decision is usually left up to the judge.

Risk-assessment-based criminal justice reform has not always had the intended impact. Judges vary widely in their propensity to release low-risk defendants. Some follow the recommendations associated with the risk assessment tool most of the time; others only rarely. As a result, risk-assessment-based reforms that are intended to lower incarceration rates may have little to no effect.

“Proponents of risk assessment have argued that they will help to lower incarceration rates with minimal effect on public safety.”
Judges are people. They come with their own beliefs about what kind of criminal justice response is appropriate in each particular circumstance. But being a judge is also a job. Understanding judicial incentives requires understanding the process by which judges are hired, retained, and promoted.

Almost all state court judges must obtain and keep their positions by means of highly-politicized election and appointment schemes. This is very different from the federal judicial system, in which, after judges are nominated by the President and confirmed by the Senate, they become largely insulated from political pressure by lifetime terms. Even after gaining office through election or appointment, almost all state court judges must routinely seek re-election or re-appointment on a regular basis. Trial court judges, who decide criminal cases, are often especially exposed to these political pressures.

Twenty-nine states elect their trial court judges, while two use a system of legislative appointment and re-appointment. Such systems, combined with relatively short terms for most trial court judges, mean that judges are routinely confronted with the reality that voters or legislators will be passing judgement on their judicial records. The incentive to avoid being characterized as “soft on crime” – and thus at risk of losing their position – is powerful.

The states that are the subject of this report – Kentucky and Virginia – employ systems of judicial selection and retention which have the potential to expose judges to significant political pressures relating to criminal justice issues. In both states the general jurisdiction trial courts with jurisdiction over felonies are called Circuit Courts. Under Kentucky law, Circuit Court judges obtain and retain office by running in nonpartisan elections, except in cases of filling an unexpired term, in which case they are appointed by the Governor. They serve eight-year terms. In Virginia, Circuit Court judges are appointed by the legislature, and also serve eight-year terms, at the conclusion of which they may be re-appointed or replaced.
Here we analyze the impacts of risk-assessment based criminal justice reforms in two different contexts: determining pretrial custody and bail, and determining sentences. We analyze the impacts of pretrial risk assessment in Kentucky, and the impacts of sentencing risk assessment in Virginia. These locations were selected for several reasons. First, their reforms occurred at the statewide level instead of at the county level. This ensures that there is a large number of observations on which to base our analysis. This also allows us to identify cross-county and cross-judge variation in compliance with the reform efforts.

Second, these states were both early adopters of algorithmic tools to aid in criminal justice decision-making. Kentucky has used some sort of algorithm to aid in the pretrial custody decision since the 1970s, and made its use mandatory in 2011. Virginia is one of the first states to incorporate algorithmic risk assessment into sentencing; they piloted risk assessment in the late 1990’s and adopted it statewide in 2002.

Virginia and Kentucky were also selected because they adopted risk assessment with the explicit goal of reducing jail and prison populations. Kentucky adopted risk assessment after a decade of substantial growth in incarceration. By 2010, the jail and prison population imposed a significant burden on the state budget. Risk assessment was proposed as a “smart on crime” method of countering this growth. Virginia adopted risk assessment after a major “truth in sentencing” reform that increased sentences for violent offenders and mandated that felony offenders would serve at least 85% of their sentence. Concerned that this reform led to over-crowded jails and prisons, the Virginia legislature adopted risk assessment as a release valve. They asked the Virginia Sentencing Commission to design a risk assessment tool that could identify 25% of the lowest risk nonviolent offenders for diversion from jail or prison.

Kentucky and Virginia thus provide interesting case studies in how risk-assessment-based criminal justice reform affects incarceration rates. While each jurisdiction will have its own unique experience, evidence from Kentucky and Virginia shows that risk-assessment-based reform may not have the intended effect.
Kentucky’s judges vary widely in the extent to which they follow the recommendations associated with the risk assessment. Some judges grant non-financial release (release without secured monetary bond) for more than 70% of low and moderate risk defendants. Others virtually never grant non-financial release. This variation can be seen in Figure 2. Each spike in the figure shows the average non-financial release rate for a particular judge. The judges are ordered from the most strict on the left to the most lenient on the right. The median judge, represented by the red vertical line, grants non-financial release to only 37% of low and moderate risk defendants. In other words, the median judge overrules the presumptive default about 2/3 of the time.

**VIRGINIA**

In 2002, Virginia adopted a risk assessment tool with the goal of identifying the 25% lowest-risk non-violent offenders for diversion from prison or jail. In this context, diversion can mean one of two things. For defendants who were recommended to be placed in prison by the sentence guidelines, diversion means either probation, community sanctions, or a shorter jail sentence. For defendants who were recommended to be placed in jail by the sentence guidelines, diversion means being placed on probation or community sanctions instead.

Figure 3: No Change in the Diversion Rate for Virginia’s Risk-Assessment Eligible Offenders as a Result of Risk Assessment
Use of the nonviolent risk assessment tool was limited to defendants convicted of a drug, larceny or fraud offense. A prior or concurrent violent conviction rendered them ineligible for risk assessment. Furthermore, since the tool was designed for diversion, it was only used on defendants who were recommended for prison or jail by the sentence guidelines. The incarceration rate also remained virtually unchanged, increasing just slightly from 81% to 82% (Figure 4).\textsuperscript{12}

As in Kentucky, Virginia judges vary widely in their propensity to follow the recommendations associated with the risk assessment. This is seen in Figure 5, in which each vertical spike shows the fraction of low-risk offenders who were granted diversion by a particular judge. The strictest judges are to the left of the graph; they divert only around 20% of low-risk offenders. The more lenient judges, shown on the right of the graph, divert almost 70%. The median judge diverts only about 40% of those who were recommended for diversion by the risk assessment instrument. (Some judges may prefer to be more lenient but feel that their jurisdiction lacks sufficient alternatives to incarceration.\textsuperscript{13} However diversion rates remain low even in many of Virginia’s most populous counties.)
There is also significant cross-county variation in adherence to the risk assessment. This is shown in Figure 6, which shows the per-county diversion rate for low-risk defendants in Virginia. There is no consistent regional pattern, nor does this diversion rate correlate strongly with population density.  

![Figure 6. Diversion Rate for Low-Risk Defendants by County](image)

**DISCUSSION**

This evidence from Kentucky and Virginia contributes to a growing literature on judicial decision making. In this section we review the prior literature and discuss how it might inform our empirical findings.

Judges come to the bench with their own preferences about sentencing, ranging from lenient to harsh in particular types of cases or with particular types of defendants. This variation is natural but undesirable in a system where we hope the law provides an objective standard and that a defendant’s outcome should not depend on which judge hears his case. While they have their own preferences, judges also respond to incentives. Judges know that some mistakes are costlier than others: It is commonly believed that voters will punish a judge for releasing someone who then commits a crime, but may not even notice when they lock up someone who is not a threat. As long as they are accountable to voters who behave in this way, judges will over-incarcerate – especially in cases where there is substantial media coverage. In addition, judges’ sentencing behavior changes at different points in the election cycle. In particular, they tend to opt for harsher sentences when they are nearing reelection, to satisfy an electorate that (surveys suggest) tends to favor harsher sentences. This variation in behavior over time is problematic because it means defendants’ outcomes depend on when their case is heard. However, it shows that voters can influence judges’ sentencing decisions when they pay attention. Community oversight will be more effective – and more fair to defendants – if it occurs throughout a judge’s term, not only during election seasons.
Legislatures have some tools at their disposal to shape judges’ behavior. The way judges are selected to begin with is important. For instance, nonpartisan elected judges and judges selected by merit commissions write higher-quality judicial opinions than partisan elected judges, as measured by future citations of those opinions. Giving commission-selected judges stronger tenure (that is, renewing their terms via uncontested retention elections) improves the quality of their opinions even more, by reducing the political pressure they face while on the bench.

Legislatively-imposed constraints on judicial discretion – such as sentencing guidelines – can also limit judges’ responses to perceived political pressure. There is only so much harsher a judge can make sentences in the run-up to an election when there is an externally-determined cap on sentence length. Similarly, mandating the use of algorithmic risk assessment scores could theoretically mitigate the worst effects of political pressure by giving judges political cover to opt for more lenient sentences when that is their preference and offenders are rated low-risk. This may be particularly beneficial when the public or media exhibits bias based on race, socioeconomic status, or gender.

As we show here, however, legislative approaches are not always effective. Both Kentucky and Virginia adopted algorithmic risk assessments with the stated goal of reducing incarceration for low-risk offenders, but incarceration rates did not fall substantially in either state. There is good reason to allow judicial discretion to override the recommendation of one-size-fits-all risk scores, sentencing guidelines, and so on. But the existence of judicial discretion means that voters cannot assume that legislative fixes will achieve their stated goals.

That’s not to say that legislative approaches like the adoption of risk assessment won’t ever result in lower incarceration rates. New Jersey adopted a comprehensive bail reform bill that included the implementation of risk assessment and jail populations plummeted. However, there are noteworthy differences between New Jersey and the states we analyze here. Among them is the fact that judges in New Jersey are appointed by the governor, and both Governor Christie and Governor Murphy have been champions of criminal justice reform. Judges face a different incentive structure when they are appointed by someone who is supportive of reform, as opposed to gaining and retaining their jobs in a general election (as is the case in Kentucky) or via a legislature that likely is not as tightly focused on carrying through an agenda (as in Virginia).

Empirical evidence shows that voters have tremendous power to shape judges’ decisions, by paying attention to their behavior on the bench and voting based on whether they like what they see. But it also shows that voters must play a continuous and active role if they want to build a criminal justice system that serves the best interests of their communities. Ultimately, policymakers interested in a less punitive and more consistent criminal justice system may also need to consider reforming how judges are selected and kept in their jobs. So long as judges are subject to politically-driven methods such as elections or legislative appointment and retention, strong incentives will remain in place to be perceived as “tough on crime.” Until judges are freed from these political pressures, expecting them to adhere to decarceral reform initiatives may be little more than wishful thinking.

“It is commonly believed that voters will punish a judge for releasing someone who then commits a crime, but may not even notice when they lock up someone who is not a threat.”
MEGAN T. STEVENSON
Assistant Professor of Law at George Mason University

Megan T. Stevenson is an economist and legal scholar who is currently an Assistant Professor of Law at George Mason University. Her research uses econometric techniques to evaluate criminal law and policy in areas such as bail, pretrial detention, juvenile justice, misdemeanors, and risk assessment. Her studies have been published in top journals in both economics and law, such as the Stanford Law Review and the Review of Economics and Statistics. Her research has been funded by the National Science Foundation, the Russell Sage Foundation, and the Laura and John Arnold Foundation. Prior to joining the law faculty at George Mason, Professor Stevenson was a fellow at the Quattrone Center for the Fair Administration of Justice at the University of Pennsylvania Law School (2015-2017). She holds a BA in Interdisciplinary Studies (2009, with highest distinction) and a PhD in Agricultural and Resource Economics (2016), both from the University of California, Berkeley. She can be found on Twitter at @MeganTStevenson.

JENNIFER L. DOLEAC
Associate Professor of Economics at Texas A&M University,
and Director of the Justice Tech Lab

Jennifer L. Doleac is an Associate Professor of Economics at Texas A&M University, and Director of the Justice Tech Lab. She is also a Nonresident Fellow in Economic Studies at the Brookings Institution. Professor Doleac studies the economics of crime and discrimination, with a particular focus on the effects of technology on public safety. Past and current work addresses topics such as DNA databases, gun violence, and prisoner reentry. Her research has been supported by several governmental and philanthropic organizations, and published in leading academic journals including the Review of Economics and Statistics, the American Economic Journal: Applied Economics, and the Economic Journal. Professor Doleac’s work has also been highlighted in an array of media outlets, including the Washington Post, the Wall Street Journal, The Atlantic, The Guardian, and Time. She holds a Ph.D. in Economics from Stanford University, and a B.A. in Mathematics and Economics from Williams College. She tweets about criminal justice research and policy at @jenniferdoleac.

The American Constitution Society works to advance our democracy by promoting the law as a force to improve the lives of all people. Progress requires both ideas and the people to carry them through positions of leadership to be realized in legal precedent, policy, and practice. ACS works to shape the future of our democracy through building a pipeline and supporting networks of progressives legal minds, shaping the debate on the issues that dominate our national discussion, and providing a platform for progressive voices to advance the law as a force that protects fundamental rights.


3 Ky. Const. § 119.


The data used for this analysis was provided by the Kentucky Administrative Office of the Courts. See Stevenson, supra note 3, for more detail.

See Ostrom & Kauder, supra note 4.

Judicial circuits that had participated in Virginia’s risk assessment pilot – and therefore had access to a nonviolent risk assessment prior to fiscal year 2003 – were not included in Figure 3 or Figure 4. Nor were they included in the diversion/incarceration statistics cited above.

The data used for this analysis was provided by the Virginia Sentencing Commission.


County identifiers are not available in the Kentucky data.


Id.