

The Effect of Risk Assessment Scores on Judicial Behavior and Defendant Outcomes

CarlyWill Sloan, George Naufal, and Heather Caspers *

Abstract

Risk assessment has been increasingly adopted in an effort to reduce pretrial detention for poor, low-risk defendants. This paper examines the impact of risk assessment using administrative data from a large Texas County. We identify effects using a regression discontinuity that exploits the overnight implementation of a risk assessment policy. Results indicate this led to a 6.5 percent increase in non-financial bond and an 8.5 percent decrease in pretrial detention, though neither effect persisted beyond two months. Additionally, the policy did not increase violent pretrial crime, though there is some suggestive evidence of increases in non-violent pretrial crime.

JEL Classification: K4, J7

*Sloan: United States Military Academy at West Point, carlywill.sloan@westpoint.edu; Naufal: Texas A&M University and IZA, gnaufal@tamu.edu; Caspers: Iowa Department of Human Rights, heather.caspers@iowa.gov
Acknowledgments: We are grateful for useful comments from Adam Bestenbostel, Dottie Carmichael, Laura Dague, Jennifer Doleac, Mark Hoekstra, Meg Ledyard, Jonathan Meer, Brittany Street, and Travis County Pretrial Services. The opinions expressed herein reflect the personal views of the authors and not those of the U.S. Army or the Department of Defense.

I. Introduction

In the United States (US), the Eighth Amendment and most state constitutions guarantee the right to non-excessive bail. However, 34 percent of all felony defendants are detained until case disposition (Bureau of Justice Statistics, 2013). Further, 90 percent of those pretrial detained are incarcerated because of their inability to post monetary bail (Bureau of Justice Statistics, 2013). In response to the overcrowding of prisons and a perception that the existing bail system disproportionately harms poor and low-risk defendants, many jurisdictions have shifted from monetary bail to a risk-based system, where defendants are released according to their risk of pretrial crime instead of their ability to pay bail or secure a bond. Supporters of risk assessment scores argue that assessing individuals based on their risk rather than income could lead to less pretrial detention, allowing low-income defendants to keep their jobs and imposing near-zero costs on the criminal justice system if defendants do not commit new crimes pretrial. Opponents claim that increasing pretrial release could increase pretrial crime due to the reduced expected penalty of future crime. It is also possible that risk assessment scores could exacerbate existing racial disparities, as these scores often include components, such as employment or criminal history, that are correlated with defendant race.¹ The purpose of this paper is to assess whether risk assessment policy reduces pretrial detention without increasing violent crime and whether it reduces conviction rates or increases racial disparities.

Although the use of risk assessment scores is rapidly expanding across the United States, there is little to no research on their causal effects on release patterns and defendant outcomes. There are two primary difficulties with estimating the effect of risk assessment score policies. First, most jurisdictions do not keep detailed records on defendants from arrest until case disposition, including recording whether they were assessed using a risk assessment score. Second, some jurisdictions only use scores for certain defendant types, often those charged with less serious crimes. Any resulting cross-sectional comparisons would be biased as those with scores are observably, and likely unobservably, different across many attributes.

We estimate the intent-to-treat effect of a risk assessment policy in Texas using data from Travis County, home to the state capital, Austin, and a large county with a population of over 1.2 million (United States Census Bureau, 2017). On January 14th, 2013, Travis County abruptly changed from not using a research-based risk assessment score at all to assigning one to nearly every inmate. Importantly, Travis County’s implementation of the risk assessment score policy was immediate, and the exact policy change date was not announced publicly. There was no slow roll-out of the policy—one day the county assigned no risk assessment scores, and the following day it assigned scores to over 80 percent of defendants. We use this sudden change to identify local effects through a regression discontinuity design. Using the timing of the policy change, we are able to compare defendants booked just before and after the policy change. The identifying assumption is that all determinants of defendant outcomes aside from the policy change vary smoothly through the policy change. Put another way, we assume that defendants who choose to commit crimes on January 12th and 13th versus January 14th and 15th are not meaningfully different except that those on the later dates received a risk assessment score. We also show empirical tests using exogenous covariates that support this assumption.

We show implementing a risk assessment score policy increases release on non-financial bonds by 6.5% and decreases pretrial detention by 8.5%. We also provide evidence that results are driven by low-income defendants and do not worsen existing racial disparities. Second, we find no effect on violent pretrial crime and can rule out meaningful increases. These results are robust to multiple inference and several robustness checks. There is some suggestive evidence that non-violent pretrial crime may increase; however, these results are not robust. In addition, we provide suggestive evidence that the risk assessment score policy reduced conviction. It is also important to note that changes from the risk assessment score are short-lived. Namely, judges returned to their old release patterns about three months after the policy.

To our knowledge, this paper provides some of the first causal evidence on the effects of pretrial risk assessment score policies. As a result, our work contributes to multiple important existing literatures. First, we contribute to a small but growing literature on risk assessment scores in

general. The majority of this literature has focused on the validity of risk assessment instruments rather than a policy's overall effect (Almond et al., 2017; Flores et al., 2017; New Jersey Courts, 2018; Meredith, Speir, and Johnson, 2007; Schmidt et al., 2017). For example, many states have validated their risk assessment score use by documenting that higher scores are correlated with higher recidivism (DeMichele et al., 2018; Latessa et al., 2010; Turner, Hess, and Jannetta, 2009; Zhang, Roberts, and Farabee, 2014). Others have focused on comparing human decisions with actuarial predictions (Chanenson and Hyatt, 2016; Dressel and Farid, 2018; Grove et al., 2000). Perhaps the most rigorous paper in this field, Kleinberg et al. (2017), used machine learning to determine what crime rates would have been if release decisions were made solely based on a risk assessment algorithm. They found that if the same number of inmates were released, but were chosen according to their algorithmic risk scores, crime rates would fall.

Conversely, there are few serious independent evaluations of risk assessment score implementation. One exception is Doleac and Stevenson (2018), which considers implementing a risk assessment score at another decision point: sentencing. This paper is most similar in spirit to Stevenson (2018a). She evaluates multiple pretrial risk assessment score policy changes in Kentucky using an event study framework. Stevenson (2018a) provided rigorous pre- and post-comparisons, concluding that the use of risk assessment scores alters bail-setting behavior and leads to increases in failures-to-appear and no decrease in pretrial crime. Importantly our risk assessment policy change is markedly different than the two used in Stevenson (2018a). Stevenson (2018a) first identifies effects for a policy change (Kentucky House Bill 463), which *required* judges to consider a previously implemented risk assessment score in bond decisions. Second, Stevenson (2018a) estimates the effect of a later policy change where Kentucky switched from its original checklist-style risk assessment, similar to the risk assessment tool used in Travis County, to a slightly more complex tool called the Public Safety Assessment. This paper is unique because it estimates the effect of a changing from no risk assessment to a checklist-style tool in an environment where judges were *not required* to use the risk assessment tool. These distinctions are important given the recent debates about and few rigorous evaluations of risk assessment

scores. Although it appears many jurisdictions may feel comfortable adopting a research-based risk assessment tool, they may not be ready to mandate its consideration in all bond decisions. Finally, we can also investigate the effects of risk assessment scores on conviction, a very significant outcome for defendants.

Our paper also relates to a number of papers on the effects of pretrial detention on defendant outcomes. In general, these papers have found that pretrial detention leads to an increased likelihood of conviction (Didwania, 2018; Dobbie, Goldin, and Yang, 2018; Heaton, Mayson, and Stevenson, 2017; Leslie and Pope, 2017; Stevenson, 2018b). Others have considered the effect of nonmonetary bail on outcomes, finding that nonmonetary bail decreases conviction rates (Gupta, Hansman, and Frenchman, 2016).

The results of this paper have important implications for policymakers. On the one hand, there is some scope for risk assessments to reduce income-based disparities and improve defendants' lives since having risk assessment scores increases non-financial bonds and decreases pretrial detention without increasing violent crime, at least in the short run. Indeed, Dobbie, Goldin, and Yang (2018) and Stevenson (2018b) show that decreases in pretrial detention are associated with greater job stability, less reliance on government assistance, and less separation from family. In addition, it is particularly promising that implementing risk assessment had these effects without generating an increase in violent pretrial crime. On the other hand, other findings provide a less optimistic view of the promise of risk assessment. First, we find suggestive evidence of an increase in non-violent pretrial crime, which suggests reducing pretrial detention may not be costless to society at large. More importantly, it is clear from this setting that effects did not persist beyond two months. This suggests that more work is necessary to understand why effects can be short-lived and what needs to be done to generate more lasting results. It also means that risk assessment, at least as implemented in this setting, is not a panacea for creating meaningful, ongoing reductions in pretrial detention.

II. Overview of the Travis County System

With a population of over 1.2 million, Travis County is one of the largest and fastest-growing counties in the nation (United States Census Bureau, 2018). It is also known as one of the first Texas counties to focus on reducing pretrial detention (Craver, 2017; Smith, 2012). In early 2013, research-based risk assessment scores were implemented by Travis County Pretrial Services for the first time. Travis County chose to implement the Ohio Risk Assessment System-Pretrial Assessment Tool (ORAS-PAT) for its risk assessment scoring. The ORAS-PAT is a relatively new risk assessment tool, developed in 2009 and validated by the University of Cincinnati. It is also one of the most commonly used risk assessment tools in the United States (Stanford Pretrial Risk Assessment Tools Factsheet Project, 2021). Although there is little work comparing risk assessment tools in the same jurisdiction, the ORAS-PAT performs similarly to other popular tools such as the PSA and VPRAI across different counties in California (Judicial Council of California, 2021).²

After a defendant is arrested and booked in Travis County, they are interviewed by a pretrial services officer. Relying on information collected during the pretrial interview and facts from a defendant's criminal history, the pretrial services officer calculates a defendant's risk assessment score. The form used by Pretrial Services to calculate a defendant's score is presented in Figure 1. Specifically, the ORAS-PAT considers age at first arrest, number of past failures to appear, prior jail incarcerations, employment status at arrest, residential stability, and drug abuse as inputs. Next, the pretrial services officer adds up the points assigned to each input, yielding a risk score. This score is used to group a defendant into one of three different categories of pretrial crime risk: low, moderate, or high. Pretrial officers often also make a recommendation to release or detain defendants pretrial based on the risk assessment score and category assigned to the defendant. If pretrial services recommends release, the recommendation is passed onto a judge. For example, if a defendant scores a one, the score will be passed along to the judge. Sometimes the pretrial services officer will also mark "recommend release" on the form as well (but not always).

Before the rollout of the risk assessment policy, judges were informed that pretrial services would use the ORAS-PAT. They also received an email explaining how Pretrial Services scores defendants and comes to a recommendation using the ORAS-PAT. However, Pretrial Services did not explicitly tell judges how to use their recommendation or the ORAS-PAT. Instead, judges likely viewed the policy change as providing additional helpful information for release decision-making.

After considering the recommendation, judges have three options at a bail hearing. First, they can award a non-financial bond, meaning the defendant is not detained pretrial and is free to return home after the hearing with no financial obligation. Second, the judge can award a financial bond, in which case the defendant must post bail (pay the amount of bail in its entirety) or pay a portion of the bail amount upfront to a bail bondsman in order to be released pretrial. In the case of financial bail, the judge does not directly determine the pretrial detention status for the defendant. Third, the judge can deny non-financial bond or financial bond, forcing the defendant to be detained pretrial.

Importantly, because judicial approval is still required for pretrial release (i.e., a defendant's bail and release decisions do not rely entirely on the recommendation from their risk assessment score), it is natural to wonder if judges even utilize risk assessment scores. While it is impossible to say definitively that all judges seriously consider risk assessment scores, 55 percent of Texas pretrial judges surveyed in Carmichael et al. (2017) stated that lack of validated risk assessment tools are a barrier to informed release decisions. Moreover, 80 percent of Texas pretrial professionals and 70 percent of judges support or do not oppose adopting pretrial risk assessment scores. Finally, according to Carmichael et al. (2017), the ORAS-PAT is considered an important source for determining non-financial bond.

After the judge's decision, a defendant is released from jail if they are awarded non-financial bond or if they pay for release,³ but they are expected to show up for all future court proceedings. If a defendant is arrested for a new crime, we say that this defendant has committed a new pretrial crime. This defendant is then likely returned to jail until the final disposition of their case. For all defendants, we say a defendant is convicted if they are found guilty by trial or accept a plea deal.

III. Data

We use individual-level administrative data from Travis County on all criminal cases disposed between 2011 and 2014. Our data come from two different sources within Travis County. First, Travis County Pretrial Services provides data on defendant characteristics, booking, risk assessment score interviews, and bond outcomes. Importantly, these also include the exact booking date for a defendant, which is essential to determining a defendant's treatment status. We combine these data with information from a second source: data on the disposition of cases and pretrial crime from the Travis County Court System.

We identify five outcomes of interest: release on non-financial bond, pretrial detention, conviction, non-violent pretrial crime and violent pretrial crime.⁴ Unfortunately, information on non-financial bonds is missing for roughly 11 percent (15,183) of defendants. Travis County Pretrial Services believes the missing data to be the result of recording oversights and is not related to the policy change or a particular type of defendant. Even so, we discuss this limitation in greater depth in Section V.F.

Non-financial bond is equal to one if a judge assigns a defendant a non-financial bond and zero for those denied bail or those assigned monetary bail. If a defendant is released on non-financial bond, they can leave jail immediately and need only promise to return to court at a later date. Pretrial detention takes on a value of one if a defendant is kept in jail for more than two days before their disposition not including time served after potential subsequent arrests.⁵ Conviction is recorded as a one if a defendant is convicted of the crime they were originally arrested for, and zero otherwise. Pretrial crime is measured for all defendants, regardless of their pretrial bond or detention status, and takes on a value of one if a defendant is arrested for a new crime before the disposition of their current case. Severity of crime is defined by the Texas Office of Court Administration. Non-violent pretrial crime takes on a value of one for all defendants who, before their trial, are arrested for a new non-violent crime. Violent pretrial crime takes on a value of one for all defendants who are arrested for a new violent pretrial crime.

Table 1 presents descriptive statistics for all defendants booked in Travis County from 2011 through 2014. Most defendants are minorities (non-white or Hispanic), male, and US citizens: 58, 76, and 89 percent respectively.⁶ Eighty eight percent of defendants are not flagged by the mental health assessment at booking. Just over half the defendants (51 percent) are also categorized as indigent. Defendants are considered indigent if they have low income, rely on certain forms of government assistance, or reside in a public mental health facility.⁷ For the entire time period, 33 percent of defendants have a risk assessment score recorded, although 77 percent of defendants have a risk assessment score after January 2013.

IV. Methods

A. Identification Strategy

For this paper, we exploit a sharp policy change that occurred in Travis County on January 14, 2013. On this date, the county fully implemented a new risk assessment score practice, shifting from not using risk assessment scores for defendants to calculating a risk assessment score for over 80 percent of defendants. This is an ideal setting for applying our regression discontinuity design to estimate the short term causal effect of a risk assessment score policy on defendant outcomes. The identifying assumption is that all determinants of defendant outcomes vary smoothly through the policy change threshold. Intuitively, we compare defendants booked just before and just after the policy change, assuming that the timing of their booking around the policy change threshold is as good as random. Given the institutional details of the policy change, it is difficult to believe that precise manipulation of the time of a crime is feasible. For manipulation to occur, a defendant must have been aware of the exact start date of the policy—which was not readily advertised to the public—and have shifted the timing of their crime accordingly. Because treatment is also determined by the defendant’s booking date and not the bail hearing date, it is unlikely that a judge would be able to alter the treatment status of a defendant. Finally, we have also spoken to many court employees, and they all communicated that no other policy changes occurred at the same

time or around the time as our risk assessment policy. Therefore, we also believe it is unlikely that our results are driven by other policy changes around the risk assessment policy.

Formally, we estimate the following individual-level OLS model following the standard Regression Discontinuity equation:

$$(1) \quad Outcome_{it} = \alpha + \beta I[policy\ enacted \geq 0]_t + \gamma days\ from\ cutoff_t + \delta I[policy\ enacted \geq 0]_t days\ from\ cutoff_t + \lambda_i + \pi_c + \rho_d + \epsilon_{it}$$

Here, i indexes individual defendants and t the date of booking. $Policy\ enacted_t$ takes on a value of one if a defendant was booked on or after the day of the policy enactment and is zero otherwise. The running variable, $days\ from\ cutoff_t$, is defined as days from the date of policy enactment, or $date\ of\ booking_t - policy\ enactment\ date_t$. By interacting $I[policy\ enacted]$ with $days\ from\ cutoff_t$, we allow the slopes of our fitted lines to differ on either side of the policy change. Our coefficient of interest, β , captures the intent-to-treat effect of the risk assessment score policy.⁸ Namely, we estimate the impact of Pretrial Services preparing, estimating, and providing judges with the risk assessment score (what we call the effect of a risk-assessment policy). This is the combined effect of offering the judges the tool and their decisions to use it. λ_i contains individual-level controls that could alter the precision of our estimates, but should not drastically change our estimates of β if our identifying assumption holds. π_c is court-specific fixed effects, and ρ_d is day-of-week fixed effects, which capture any time-invariant court tendencies or differences across days of the week, respectively. Finally, the error term, ϵ_{it} , measures any unobservable factors that could also alter outcomes.

Our preferred specification employs the mean square error (MSE) optimal bandwidth suggested by Calonico et al. (2017). As is standard in the regression discontinuity literature, we report results for various other bandwidths and show that our main results are not sensitive to bandwidth choice. Finally, we report robust standard errors calculated as suggested in Calonico et al. (2017).⁹

B. Tests of Identification

Given the nature of the policy change noted before and the late implementation of the policy (i.e., not on January 1), we believe it to be unlikely that defendants or judges could have manipulated the assignment of treatment in a manner that would discredit our research design. Even so, we provide empirical evidence that our identifying assumption is valid by demonstrating that the number of defendants booked, as well as observable defendant and case characteristics, do not vary discontinuously through the policy change threshold. Figure 2 shows the distribution of the running variable, days from the cutoff. If manipulation were possible, we would expect to see a spike or fall in the number of defendants booked, but this is not the case.

Next we investigate if specific case and defendant characteristics are smooth through the policy change threshold. If our identifying assumption is valid, defendant and case characteristics will vary similarly on both sides of the policy change threshold. If defendants or judges could have exactly manipulated the timing of booking, we would expect to find differences in case and defendant characteristics through the policy enactment threshold. To test this threat to identification, we estimate Equation (1) using race, age, gender, criminal history, indigent status, severity of arrest (misdemeanor or felony), mental health status, US citizenship status, and specific court separately as outcome variables. Figure 3 and Table A.1 (Online Appendix) show the results for this test. There are no visible jumps in defendant and case characteristics in the graphs presented. Of the 20 estimates presented in Table A.1 (Online Appendix), only two are statistically significant at conventional levels, although with coefficients close to zero, which is consistent with findings due to chance. These results indicate that case and defendant characteristics are not discontinuous through the policy change threshold.

We also present another test of the identifying assumption using all the covariates we observe about a defendant and case that are determined before defendants are assigned a risk assessment score. Instead of considering the covariates individually, we use them in combination along with a court and day-of-week fixed effect to predict the likelihood of each potential outcome (release on

non-financial bond, pretrial detention, non-violent pretrial crime, violent pretrial crime, and conviction) for every defendant. This allows us to create a weighted average where the characteristics that contribute more to a specific outcome are considered with greater weight. Here we can estimate the underlying probability of an outcome using everything we know about them except the use of a risk assessment score. If each predicted outcome is smooth through the policy change threshold, then we can attribute any treatment effect we later estimate to the policy change, not underlying differences in defendants booked just before and after the policy change.

Figure 4 and Table A.2 (Online Appendix) show the results for the predicted outcomes. The regression discontinuity estimates for each predicted outcome are statistically insignificant and are close to zero. This further indicates little evidence of underlying differences in defendants across the policy change threshold—proving further that our identifying assumption holds.

V. Results

A. Effects of Risk Assessment Score Policy on Score Usage

To determine the effects of a risk assessment score policy, we first need to document that Travis County’s enactment of its risk assessment score policy led to a sudden and dramatic increase in the number of defendants assessed and assigned a risk assessment score. To do so, we estimate Equation (1) using the assignment of a risk assessment score as the outcome variable. Figure 5 presents our graphical results. This graph and the graphs that follow plot the mean of the outcome variable in two week bins. In all figures, the running variable is normalized to zero (the date of policy enactment is zero days after the policy change).

Figure 5 shows clearly that we estimate a large (about 80 percent) increase in risk assessment score assignment across the policy change threshold.¹⁰ This indicates that about 80 percent of defendants booked after the policy enactment were assigned a risk assessment score. We note that this is not a sharp discontinuity (i.e., 100 percent take-up), which motivates our use of intent-to-

treat estimates throughout the rest of the paper.¹¹ There are multiple reasons why a defendant may not have been recorded with a risk assessment score. First, the Pretrial Services data we use are not perfect. It may be the case that some scores simply were not recorded. Furthermore, some defendants are much less likely to receive a risk assessment score, such as those who have an active defense attorney to convince Pretrial Services not to conduct a pretrial risk assessment score or those with a parole violation. Regardless, our intent-to-treat effects allow us to estimate the unbiased intent-to-treat effect of the risk assessment score policy.

B. Effects of a Risk Assessment Score Policy on Non-financial Bond and Pretrial Release

The primary intent of the risk assessment adoption was to increase the number of defendants released on non-financial bond. This decision is made by judges with access to risk assessment scores, so this is the first outcome we consider. Next, we consider pretrial detention. If a defendant is released on non-financial bond, they are not detained pretrial; but if a judge offers financial bond to a defendant, their pretrial detention status is determined by their ability to pay the bond. Therefore, it is of separate interest to determine the effects of a risk assessment score on pretrial detention.

We first show the effects of a risk assessment score on non-financial bonds and pretrial detention in Figure 6. Formally, we estimate Equation (1) with the probability of release on non-financial bond and pretrial detention as outcome variables. Figure 6 shows the mean of release on non-financial bond and pretrial detention in two week bins. This figure provides visual evidence that implementing risk assessment scores increases the likelihood of release on non-financial bond and decreases pretrial detention in the first two months after the policy change. It also appears these effects fade with time. While it is challenging to determine exactly why our results decrease with time, we will discuss possible reasons later in this section.

Table 2 presents corresponding point estimates, with each column representing a separate

regression. Column 1 presents our baseline regression discontinuity results. Column 2 include controls for days since the policy change, case specific controls, and fixed effects for the court and day of booking. Specifically each column has case-level controls for defendant race, age, gender, citizenship, mental health flag and indigent status, along with controls for the severity of the crime (misdemeanor or not). Each specification allows the running variable, days from the policy enactment, to vary linearly on each side of the cutoff. Both columns show results for the optimal bandwidth. If our identifying assumption holds, we would expect that our coefficient of interest would remain similar in magnitude. Across both columns, our estimates remain statistically significant at conventional levels and are of similar magnitudes for non-financial bond and pretrial detention.¹²

Our estimates for non-financial bond range from 0.0419 to 0.0430. These results indicate that the implementation of a risk assessment score policy increases the likelihood of release on non-financial bond by about 4 percentage points (6-6.5%). For pretrial detention, our estimates range from -0.0273 to -0.0303, showing the risk assessment score policy decreases the chance of pretrial detention by about 3 percentage points (8.5-9.5%).

Because our results include five different outcomes, we also include false discovery rate (FDR)–adjusted q-values for the estimates presented in Table 2. We compute the FDR-adjusted q-values using the method proposed by Anderson (2008), adjusting for our five different outcomes. The FDR q-values can be interpreted as adjusted p-values. The FDR q-values for each outcome are statistically significant at conventional levels for each specification. Therefore, we conclude that the effects we find are large enough not to be attributable to chance.

Now we will demonstrate that our results for pretrial detention and non-financial bond are robust to various specifications. A standard concern with regression discontinuity estimates is that results are valid only for a specific bandwidth selection or are the result of misfitting the data. To address these concerns, we present several specifications and show that our results are robust to bandwidth selections. First, we estimate Equation (1) with inclusion of the controls and allow the bandwidth to vary from 20 to 660 days in 10-day increments using a linear specification. Figure

8 Panels (a) and (b) show the coefficients and standard errors from each model for non-financial bond and pretrial detention. The dashed lines represent the optimal MSE bandwidth. The estimated coefficients remain consistent across the different bandwidths. Estimates for non-financial bond and pretrial detention are also statistically significant for the vast majority of estimates, illustrating that our results are robust to alternative specifications of bandwidth.

We also conduct a permutation test in the spirit of Abadie, Diamond, and Hainmueller (2010) to support our claim that Travis County’s risk assessment score policy drives our results. This test also addresses a specific concern about the time-series nature of our data. Specifically, that our errors terms are serially correlated, potentially leading to incorrect standard errors. To do so, we estimate Equation (1) reassigning the policy threshold to be a day before the true policy change occurred. Because we only have data beginning in 2011, we estimate Equation (1) 910 times using every possible date that occurred before the true policy change, a linear specification, optimal bandwidth, and the controls included in Table 2. The distribution of placebo estimates for release on non-financial bond and pretrial detention are shown in Figure A.1 (Online Appendix). Nearly all placebo coefficients (98%) are less than the reported estimates in Table 2 for release on non-financial bond. Our pretrial detention estimate in Table 2 is less than 95 percent of our placebo estimates. These simulations imply our estimates are not simply due to chance.

Finally, we consider whether our results are driven by seasonal patterns in non-financial release or pretrial detention. This is of particular concern because our running variable is time and it is possible that our results are due to annual fluctuations in cases brought to the court. To address this concern we consider whether we see our treatment effect in other years. In Figure 9 we show our standard regression discontinuity graphs for the true treatment year and the other years in our data (2011, 2012 and 2014). If it were the case that our result could fully be explained by seasonality we would expect the plots for our year of treatment and 2011/2012/2014 to be very similar before and after the policy change date (January 14 each year). In the days before the policy change the black and gray dots are very similar to each other. However, just after the policy change it is clear that the black dots (our treatment year data) are visibly different from the other years when treatment

did not occur. In Figure 9 (a) the mean of non-financial bond for the true treatment year (black dots) is visibly higher than the gray dots (2011/2012/2014). In Figure 9 (b) the treatment year dots (black dots) are visibly lower than the gray dots (2011/2012/2014).

In Table A.3 (Online Appendix), we show the results using the placebo treatment years in 2011, 2012 and 2014 for our release results. Each cell is a separate regression using optimal bandwidth. Column 2 includes controls. If our results are due to seasonal changes that occur annually, then we should expect to replicate our results in the three years we analyze. However, every estimate in Table A.3 (Online Appendix) is smaller in magnitude than the estimates in Table 2. Further no estimate is statistically significant at traditional levels (compared to all four estimates in Table 2). Our results in Figure 9 and Table A.3 (Online Appendix) show that our treatment effect is unique to 2013 and not driven by seasonality.

C. Effects of a Risk Assessment Score Policy on Pretrial Crime and Conviction

If it is the case that the new type of individuals released pretrial through non-financial bond disproportionately commit crimes before their trial, there would be an increase in pretrial crime. Results for non-violent and violent pretrial crime are shown in Figure 7. All graphs in Figure 7 show the mean of the outcome variable in two week bins. Figure 7(a) and Figure 7(b) presents some suggestive evidence of a small increase in non-violent recidivism and no change in violent recidivism. Table 3 presents the corresponding estimates. Similar to Table 2, even columns add controls for defendant race, age, gender, citizenship, mental health status, and indigent status, as well as the severity of the crime (misdemeanor or not). Each specification allows the running variable to vary linearly. Fixed effects for the assigned court and booking day of the week are also included. Importantly, our estimates for non-violent and violent pretrial crime are of similar magnitudes in the two columns. For non-violent pretrial crime, estimates range from 0.01 to 0.0117 (9-10.5%). Neither estimate is significant at traditional levels.

Next we consider violent pretrial crime. Across both columns our estimates are similar (-0.002 and -0.003). We are also able to rule out any increase in violent pretrial crime for most specifications with larger bandwidths (see Figure 8). We also report FDR q-values for pretrial crime outcomes.¹³ The FDR q-values are also not consistently significant for any outcome.

We can also show our pretrial crime results are robust to alternative bandwidths. As we did for non-financial bond and pretrial detention, we estimate Equation (1) using non-violent and violent pretrial crime as outcomes, with the inclusion of the controls and allow the bandwidth to vary from 20 to 660 days in 10-day increments. Figure 8 shows the coefficients and standard errors from each model for non-violent and violent pretrial crime in Panels (c) and (d). The dashed lines represent the optimal MSE bandwidth. The estimated coefficients remain consistent across the different bandwidths for violent pretrial crime. However, results are noisy at small bandwidths for non-violent pretrial crime. Together these results indicate that risk assessment scores do not increase violent pretrial crime. We also find some evidence, although not robust, of increases in nonviolent pretrial crime.

Finally, we consider conviction. Given the existing literature on the effects of pretrial detention on conviction, risk assessment scores could also alter conviction (Dobbie, Goldin, and Yang, 2018; Stevenson, 2018b). Figure 7(c) shows our results for conviction. Here we show some evidence of a decrease in conviction after the adoption of a risk assessment score. Corresponding estimates are found in Table 3. Both coefficients are negative and show that risk assessment scores decreased the odds of conviction for defendants. The magnitude of the coefficient remains similar across specifications, ranging from -0.0203 to -0.016, translating to decreases in conviction of 3-4%. However, Column 1 is not significant at conventional levels and neither FDR q-value is significant at the 5% level. The distribution of placebo estimates for conviction is shown in Figure A.1 (Online Appendix). In line with the p-values for our estimates in Table 3, our placebo estimate shows 11% percent of estimates are smaller than our original. Figure 8(e) shows consistent results for different bandwidths. Together, these results show potential decreases in conviction for defendants, which is consistent with findings in Dobbie, Goldin, and Yang (2018) and Stevenson (2018b) that pretrial

detention increases conviction rates.

D. Indigent Defendants

Since one stated aim of the risk assessment score policy was to improve outcomes for low-income defendants, we also present results for indigent versus non-indigent defendants. As indigent defendants were more likely to be unable to post their bond before the policy change, we would expect effects for release on non-financial bond and pretrial detention to be stronger for indigent defendants compared to non-indigent defendants. Our graphical results are shown in Figure A.2 (Online Appendix). Entire sample results are replicated in Panel (a) and (d) for release on non-financial bond and pretrial detention. Panels (b) and (e) present results for indigent defendants, while Panels (c) and (f) show results for non-indigent defendants. For both outcomes the discontinuity for indigent defendants is visibly larger than for non-indigent defendants. There is also some evidence of an increase in release on non-financial bond and a small decrease in pretrial detention for non-indigent defendants. Corresponding estimates are shown in Table 4.

In Table 4, Panel A presents results for non-indigent defendants and Panel B shows results for indigent defendants. Similar to earlier result tables, even columns include controls. All columns are a linear specification. Across each specification the coefficient for non-financial bond and pretrial detention for indigent defendants has a greater magnitude, roughly two to three times larger, than for non-indigent defendants, although we cannot rule out that the estimates are statistically equivalent. As we would expect given the aim of the policy, these subgroup results suggest that our release on non-financial bond and pretrial detention results are likely driven by low-income defendants.

We also explore results by indigent status for pretrial crime. As indigent defendants are the most likely to be released pretrial, it is possible that changes in their pretrial crime behavior are masked in the entire sample results. Results for non-violent and violent pretrial crime are shown in Figure A.3 (Online Appendix). Panels (a), (d) and (g) repeat the entire sample results for comparison. Indigent results are shown in Panels (b), (e) and (h), while Panels (c), (f) and (i)

report results for non-indigent defendants. For non-violent pretrial crime, there is some evidence of a larger increase in pretrial crime for indigent defendants and no increase for non-indigent defendants. For violent pretrial crime, however, there appears to be no increase for indigent or non-indigent defendants. Conviction results also appear similar for indigent and non-indigent defendants.

Table 5 shows pretrial crime and conviction estimates. For non-violent pretrial crime, the coefficients for indigent defendants are larger in magnitude than for non-indigent defendants. For violent pretrial crime there are no meaningful differences in the coefficients for indigent and non-indigent defendants. For each subgroup the coefficient for violent pretrial crime is negative, again suggesting there are no increases in violent pretrial crime across either group. Finally, for conviction the results are a bit larger for non-indigent defendants.

In summary, our results for indigent versus non-indigent defendants show that lower income defendants are the most likely to be awarded non-financial bond and released pretrial. We also find some suggestive evidence that non-violent pretrial crime may increase for lower income defendants, who are most likely to be released. Violent pretrial crime does not increase for either group and estimates suggest decreases in conviction for both groups.

E. Minority Defendants

There is growing concern that risk assessment scores may exacerbate existing racial disparities. For example, before the adoption of risk assessment scores, minority defendants were 10 percent less likely to be released on non-financial bond and 40 percent more likely to be pretrial detained than non-minorities.¹⁴ Components of the ORAS-PAT such as employment, prior incarcerations and drug use may be correlated with race. This means that even though the ORAS-PAT does not directly consider race, it could lead to different outcomes for minority versus white defendants. Graphical non-financial bond and pretrial detention results for white and minority defendants are shown in Figure A.4 (Online Appendix). Entire sample results are replicated in Panel (a) and (d) for release on non-financial bond and pretrial detention. Panels (b) and (e) present results for white

defendants, while Panels (c) and (f) show results for minority defendants. For both outcomes the discontinuity for minorities is potentially larger than for white defendants. Corresponding estimates are shown in Table 6. For non-financial bond and pretrial detention, in each specification, the magnitude of our estimate is larger (nearly double in each specification for non-financial bond) for minority defendants than white defendants.

Results for non-violent and violent pretrial crime and conviction are shown in Figure A.5 (Online Appendix). Panels (a), (d), and (g) repeat the entire sample results for comparison. White defendant results are shown in Panels (b), (e) and (h) while Panels (c) (f) and (i) report results for minority defendants. For non-violent crime, there is some evidence of a larger increase in pretrial crime for minority defendants and no increase for white defendants. For violent pretrial crime and conviction, however, there appears to be similar changes for white and minority defendants.

Table 7 shows pretrial crime and conviction estimates. For non-violent pretrial crime, the coefficients for minority defendants are larger in magnitude than for white defendants. For violent pretrial crime, however, there are no meaningful differences in the coefficients for minority and white defendants. For each subgroup the coefficient for violent pretrial crime is negative, again suggesting there are no increases in violent pretrial crime across either group. Finally, conviction results appear to be driven by white defendants.

Taken together these results show that our results for non-financial bond and pretrial detention are in part driven by minority defendants. There is no evidence that the adoption of risk assessment scores increased pre-existing racial disparities in non-financial bond or pretrial detention. If anything, the magnitude of our coefficients suggests that risk assessment score may have decreased racial disparities at least for non-financial bonds and pretrial detention.

F. Missing Values

One limitation of this study is that we are missing one outcome variable (non-financial bond) for 10 percent of our defendants. Although our institutional details, namely that Travis County Pretrial Services believes that some records are simply missing by chance, indicate that missing outcomes

is not correlated with treatment, we also provide empirical evidence that the likelihood of missing the probability of release on non-financial bond is not discontinuous through the threshold. Results of this test are shown in Figure A.6 (Online Appendix). Here we estimate Equation (1) using the probability of missing data for release on non-financial bond as the outcome variable. There is no striking visual evidence that the probability of missing data changes through the policy change threshold.

We provide corresponding point estimates in Table A.4 (Online Appendix). Columns (1)-(2) use the optimal bandwidth determined in Table 2 for release on non-financial bond. Columns (3)-(4) use the optimal bandwidth for the probability of missing data. Across all four columns, the coefficient remains close to zero and no estimate is statistically significant at conventional levels. Together these results indicate that the probability of missing data does not vary with treatment.

One might remain concerned that there are changes in the composition of defendants' missing data that coincide with treatment. For example, it could be the case that we are missing data for defendants who are likely to be released on non-financial bond to the left of the threshold and are missing data for defendants who are not likely to be released on non-financial bond to the right of the threshold. Although we cannot assess this directly, we can use the case and defendant information we do observe about all defendants to predict the likelihood of release on non-financial bond for defendants who are missing this outcome. We then estimate Equation (1) using predicted probability of release on non-financial bond as the outcome just for defendants who are missing data. Figure A.7 (Online Appendix) shows these results. There is no visual evidence of underlying differences in defendants who are missing data across the threshold. Taken together, these results indicate that it is unlikely that the defendants with outcomes missing from our dataset are sufficiently different to alter our results for release on non-financial bond.

G. Long Term Effects and Seasonality

We now turn to the reasons why we only observe short-term effects that fade over time. Because our regression discontinuity estimates only allow us to obtain local average treatment effects—or,

in other words, we can only establish the causal effects of the risk assessment score policy just around the time of the policy change—we cannot credibly identify long-term effects of the policy change. However, we can provide suggestive evidence related to the timing of when the effects of the policy begin to fade. To do so, we first conduct an event study analysis with results presented in Figure 10 (a) and (c). To investigate how our effect changes over time, we need to compute a counterfactual for what release decisions would have been in the absence of the risk assessment policy change. To do so, we compare differences in non-financial release and pretrial detention in the six months after the true policy change to the six months before, relative to the difference in the same time periods in 2011, 2012, and 2014. Specifically, Figure 10 (a) and (c) show coefficients from the regression of non-financial bond, and pretrial detention on indicators for months before or after risk assessment adoption interacted with an indicator for our true treatment time period. We also include fixed effects for months before/after “treatment” and each time period.¹⁵ Individual level controls for race, age, gender, citizenship, and indigent status of the defendant along with controls for the severity of the crime (misdemeanor or not) as well as fixed effects for the court assigned and day-of-week of booking are used.

Additionally, Figure 10 (b), and (d) show the time trend for each outcome estimated using local linear smoothing with a bandwidth of 14 days and controlling for seasonality using week-of-the-year fixed effects). The x-axis indicates the booking date. Pink dashed lines mark dates of the policy change in other years. The red dashed line marks the true policy change.

Visually, it is evident in all four subfigures that the effect of risk assessment scores is short-term (about two months according to the event studies) and that the rate of release on non-financial bonds, and pretrial detention return to previous levels afterward¹⁶. Further, these figures also provide more evidence that our results are not explained by seasonality. In all the subfigures, there is a fundamentally different jump just after January 14 in our treatment year compared to 2011, 2012, and 2014.

It is natural to wonder why we see such a short-lived effect from the risk assessment scores. First, we note that according to Travis County Pretrial Services, judges and pretrial service

employees did receive training on the ORAS-PAT near its implementation and that potential enthusiasm surrounding the policy could have led to short-term effects. For example, judges could have paid closer attention to the scores right after the training but stopped as time passed. It is also possible that judges began to disregard the scores after the novelty of the policy change wore off. Lastly, we note that the temporary effect documented here is consistent with evidence from Stevenson (2018a), who also provided evidence that the effect of risk assessment scores fades with time.

To address whether average treatment effect estimates mask important heterogeneities that could explain why our results fade over time, we also estimate the event studies for important subgroups in Figure A.8 (Online Appendix, non-financial bond) and Figure A.9 (Online Appendix, pretrial). There is no subgroup for which the impact of risk assessments is long term. There are some figures where no strong treatment effect is visible (felonies, white defendants, and non-indigent defendants). The other figures show only short-term effects. These results align with our findings for indigent and minority defendants in Section V.D and V.E. Further, these figures are consistent with a behavioral response by judges or pretrial service employees that is similar across different types of defendants and cases.

Regardless of why the results diminish with time, the short-term nature of effects highlights an important aspect of risk assessment scores. In practice most risk assessment scores are implemented within a pre-existing pretrial system and judges are not required to adhere to their recommendation. Inherently, any effect risk assessment scores could have on outcomes depends on how judges and pretrial services use them in their decision making process. Policy-makers must be careful to consider not only if they want to implement risk assessment scores, but also how they will be used in practice.

VI. Conclusion

This paper estimates the effects of a risk assessment score policy by using a regression discontinuity design. We compare defendants booked barely before and after a policy change in a large county in Texas. Our results indicate that implementing a risk assessment score policy leads to an increased likelihood of release on non-financial bond and a decreased probability of pretrial detention. Precisely, we estimate that the implementation of risk assessment scores led to an 6.5% increase in non-financial bonds and a 8.5% decrease in pretrial detention. We also find no increases in violent pretrial crime. We recognize that our results are only for one county in Texas and that the extent to which they apply to other contexts outside of Texas, where existing pretrial systems may be different, is unknown. Further, our results appear to be short-lived. Even with these qualifications, we believe that this study is an important contribution to nearly nonexistent literature on risk assessment scores in practice. Our results indicate that risk assessment scores have the potential to decrease costs to society and the disproportionate burden of financial bail for low-income defendants, while not increasing violent pretrial crime or racial disparities. However, policy makers must be careful to weigh these potential benefits against the chance of increases in non-violent pretrial crime. Finally, and perhaps most importantly, our results indicate that risk assessment score policies are unlikely to be the panacea that reform advocates are hoping for.

Notes

¹Academics and journalists express varying degrees of concern about potential racial biases in risk assessment scores (e.g., Angwin et al., 2016; Doleac and Stevenson, 2017).

²See Tables 8, 28, and 16, which assess the accuracy of these tools using Area Under the Curve. Intuitively, the Area Under the Curve value is a single number representing the tool's ability to distinguish between low or high risk individuals correctly (higher values indicate a better tool).

³A defendant can either pay their bail in full or a bail bondsman can post bail instead.

⁴It might also be natural to consider failure to appear. Unfortunately, Travis County does not keep accurate data on this outcome. In fact, in most Texas counties, failure to appear is not measured accurately or tracked.

⁵Our results are similar in significance and magnitude if we define pretrial detention as being in jail for more than one or three days.

⁶Travis County records the race and ethnicity of each defendant. Defendants are white if they are white and not Hispanic. Minority defendants are either non-white or Hispanic.

⁷This is the definition of indigence from Travis County Criminal Courts (2012).

⁸Roughly 80% of defendants are assigned a risk assessment score and as researchers we cannot observe whether or not a judge meaningfully considered the risk assessment score. Therefore we refer to our estimates as intent-to-treat.

⁹Although we do not believe release decisions or pretrial crime should be correlated for defendants booked on the same day, we also estimate our results clustering on booking date. For non-financial bond and pretrial detention our results have similar significance. Specifically, for non-financial bond the significance level is the same for both estimates in Table 2. For pretrial

detention, column (1) is statistically significant at the 10% level and column (2) is still statistically significant at the 1% level.

¹⁰Risk assessment usage was greater than zero for a few months before January 2013 because Travis County elected to run a pilot study.

¹¹We define our treatment as the policy change, not score usage by judges (which is unobserved). However, it is also possible to think of our estimates as intent-to-treat because it is unlikely all judges will use the score.

¹²Our results are similar in significance and magnitude if we define pretrial detention as being in jail for more than one or three days.

¹³Again, we correct for 5 categories.

¹⁴In this setting minority defendants are non-white and Hispanic. White defendants are only white.

¹⁵For example, the twelve-month period that includes January 2012, a “fake” treatment, would be considered a time period.

¹⁶Our coefficients in the two months after the policy change are statistically different (5% or less) than the coefficient in the month before the policy change in each specification.

References

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490): 493–505.
- Almond, Louise, Michelle McManus, David Brian, and Daniel Peter Merrington. 2017. "Exploration of the Risk Factors Contained within the UK's Existing Domestic Abuse Risk Assessment Tool (DASH): Do These Risk Factors Have Individual Predictive Validity Regarding Recidivism?" *Journal of Aggression, Conflict and Peace Research* 9(1): 58–68.
- Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484): 1481–1495.
- Angwin, Julia, Jeff Larson, Lauren Kirchner, and Surya Mattu. 2016. "Machine Bias." *ProPublica*, May 23. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (accessed on 01/31/2001).
- Bureau of Justice Statistics. 2013. "Felony Defendants in Large Urban Counties 2009 Statistical Tables." *Bureau of Justice Statistics (BJS)*.
<http://www.bjs.gov/index.cfm?ty=pbdetail&iid=4845> (accessed on 01/31/2001).
- Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik. 2017. "Rdrobust: Software for Regression Discontinuity Designs." *Stata Journal* 17(2): 372–404.
- Carmichael, Dottie, George Naufal, Steve Wood, Heather Caspers, and Miner Marchbanks. 2017. "Liberty and Justice: Pretrial Practices in Texas." Texas A&M University, Public Policy Research Institute.
- Chanenson, Steven L. and Jordan M. Hyatt. 2016. "The Use of Risk Assessment at Sentencing: Implications for Research and Policy." Villanova Law, Public Policy Research Paper.

- Craver, Jack. 2017. "Travis County: No Place for Bondsmen." *Austin Monitor*, March 30.
<https://www.austinmonitor.com/stories/2017/03/travis-county-no-place-bondsmen/> (accessed on 01/31/2001).
- DeMichele, Matthew, Peter Baumgartner, Michael Wenger, Kelle Barrick, Megan Comfort, and Shilpi Misra. 2018. "The Public Safety Assessment: A Re-validation and Assessment of Predictive Utility and Differential Prediction by Race and Gender in Kentucky." Social Science Research Network.
- Didwania, Stephanie Holmes. 2018. "The Immediate Consequences of Pretrial Detention: Evidence from Federal Criminal Cases." Working Paper.
- 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.
- Doleac, Jennifer L. and Megan Stevenson. 2017. "Are Criminal Risk Assessment Scores Racist?" *Brookings*, August 22.
<https://www.brookings.edu/blog/up-front/2016/08/22/are-criminal-risk-assessment-scores-racist/#7s8d6f87> (accessed on 01/31/2001).
- _____. 2018. "Algorithmic Risk Assessment in the Hands of Humans." Working paper.
- Dressel, Julia and Hany Farid. 2018. "The Accuracy, Fairness, and Limits of Predicting Recidivism." *Science Advances* 4(1): eaao5580.
- Flores, Anthony W., Alexander M. Holsinger, Christopher T. Lowenkamp, and Thomas H. Cohen. 2017. "Time-free Effects in Predicting Recidivism Using Both Fixed and Variable Follow-up Periods: Do Different Methods Produce Different Results." *Criminal Justice and Behavior* 44(1): 121–137.
- Grove, William M., David H. Zald, Boyd S. Lebow, Beth E. Snitz, and Chad Nelson. 2000. "Clinical Versus Mechanical Prediction: A Meta-analysis." *Psychological Assessment* 12(1): 19.

- Gupta, Arpit, Christopher Hansman, and Ethan Frenchman. 2016. “The Heavy Costs of High Bail: Evidence from Judge Randomization.” *The Journal of Legal Studies* 45(2): 471–505.
- Heaton, Paul, Sandra Mayson, and Megan Stevenson. 2017. “The Downstream Consequences of Misdemeanor Pretrial Detention.” *Stanford Law Review* 69(3): 711.
- Judicial Council of California. 2021. “Pretrial Risk Assessment Tool Validation.” https://www.courts.ca.gov/documents/Pretrial_Risk_Assessment_Tool_Validation_June_2021.pdf (accessed on 01/31/2001).
- 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.
- Latessa, Edward J., Richard Lemke, Matthew Makarios, and Paula Smith. 2010. “The Creation and Validation of the Ohio Risk Assessment System (ORAS).” *Federal Probation Journal* 74: 16.
- Leslie, Emily and Nolan G. Pope. 2017. “The Unintended Impact of Pretrial Detention on Case Outcomes: Evidence from New York City Arraignments.” *The Journal of Law and Economics* 60(3): 529–557.
- Meredith, Tammy, John C. Speir, and Sharon Johnson. 2007. “Developing and Implementing Automated Risk Assessments in Parole.” *Justice Research and Policy* 9(1): 1–24.
- New Jersey Courts. 2018. “New Jersey Court Report to the Governor and the Legislature.” New Jersey Court Report. <https://njcourts.gov/courts/assets/criminal/2018cjrannual.pdf?c=tap> (accessed on 01/31/2001).
- Schmidt, Nicole, Emily Lien, MaryFaith Vaughan, and Matthew T. Huss. 2017. “An Examination of Individual Differences and Factor Structure on the LS/CMI: Does This Popular Risk Assessment Tool Measure Up?” *Deviant Behavior* 38(3): 306–317.
- Smith, Jordan. 2012. “Keeping People in Jail Costs the County Money, But Is It in the Best Interest of Public Safety?” *The Austin Chronicle*, 12 October.

<https://www.austinchronicle.com/news/2012-10-12/your-word-is-your-bond/all/> (accessed on 01/31/2001).

Stanford Pretrial Risk Assessment Tools Factsheet Project. 2021. “Risk Assessment Factsheet.”

Stevenson, Megan T. 2018a. “Assessing Risk Assessment in Action.” *Minnesota Law Review* 58.

_____. 2018b. “Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes.” *Journal of Law, Economics and Organization* 34(4): 511–542.

Travis County Criminal Courts. 2012. “Travis County Criminal Courts Fair Defense Act Program.” <http://tidc.tamu.edu/DGReportDocuments/212-07-D08%20Final4.07Fair%20Defense%20Act.pdf> (accessed on 01/31/2001).

Turner, Susan, James Hess, and Jesse Jannetta. 2009. “Development of the California Static Risk Assessment Instrument (CSRA).” Center for Evidence-Based Corrections Working Paper, UC Irvine, Irvine, CA.

United States Census Bureau. 2017. “U.S. Census Bureau Quickfacts: Travis County.” <https://www.census.gov/quickfacts/fact/table/US/PST045217> (accessed on 01/31/2001).

_____. 2018. “County Population Totals and Components of Change: 2010-2017.” <https://www.census.gov/data/datasets/2017/demo/popest/counties-total.html> (accessed on 01/31/2001).

Zhang, Sheldon X., Robert E. L. Roberts, and David Farabee. 2014. “An Analysis of Prisoner Reentry and Parole Risk using COMPAS and Traditional Criminal History Measures.” *Crime & Delinquency* 60(2): 167–192.

Table 1
Summary Statistics

	Mean	Standard Deviation	Number of Observations
Case and Defendant Characteristics			
White Defendant	0.42	0.49	143,077
Minority Defendant	0.5766	0.494	143,077
Misdemeanor	0.6772	0.468	143,077
Defendant Age	32.5307	11.220	143,077
US Citizen	0.8941	0.308	143,077
Male	0.7578	0.428	143,077
Indigent	0.5124	0.500	143,077
No Priors	0.7638	0.425	143,077
Mental Health Flag	0.1238	0.329	143,077
Outcomes			
Non-financial Release	0.63	0.48	127,894
Pretrial Detention	0.3532	0.478	143,077
Violent Pretrial Crime	0.0154	0.123	143,077
Non-violent Pretrial Crime	0.1056	0.307	143,077
Conviction	0.4904	0.500	143,077

Notes: Each observation is a separate case. Data are from Travis County Courts and Travis County Pretrial Services for the years 2011-2014. Travis County records the race and ethnicity of each defendant. Minority defendants are either non-white or Hispanic

Table 2

Release Regression Discontinuity Results

	Optimal Bandwidth	
	(1)	(2)
Outcome: Non-financial Bond		
RD Estimate	0.0419**	0.0430***
	(0.0163)	(0.0143)
Observations	12572	12572
Outcome Mean	0.663	0.663
FDR q-value	0.052	0.014
Bandwidth	92.78	92.51
Outcome: Pretrial Detention		
RD Estimate	-0.0303**	-0.0273***
	(0.0135)	(0.0103)
Observations	17992	18862
Outcome Mean	0.320	0.321
FDR q-value	0.064	0.021
Bandwidth	117.5	123.7
Controls	-	Y
Running Variable Control	Y	Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Each cell shows results for a separate regression. Each panel shows results for a different dependent variable and the key independent variable is an indicator for policy enactment. Robust standard errors are in parentheses. All specifications control for the distance from policy enactment. The optimal (MSE) bandwidth is used to determine the sample for each separate regression. We compute the FDR-adjusted q-values using the method proposed by Anderson (2008), adjusting for our five different outcomes.

Table 3

Pretrial Crime and Conviction Regression Discontinuity Results

	Optimal Bandwidth	
	(1)	(2)
Outcome: Non-Violent Pretrial Crime		
RD_Estimate	0.0117	0.0101
	(0.00762)	(0.00720)
Observations	26456	28924
Outcome Mean	0.109	0.109
FDR q-value	0.432	0.342
Bandwidth	170.3	186.0
Outcome: Violent Pretrial Crime		
RD_Estimate	-0.00229	-0.00305
	(0.00291)	(0.00321)
Observations	29454	23258
Outcome Mean	0.0154	0.0153
FDR q-value	0.207	0.203
Bandwidth	189.5	150.3
Outcome: Conviction		
RD_Estimate	-0.0161	-0.0203*
	(0.0131)	(0.0107)
Observations	22494	29616
Outcome Mean	0.480	0.480
FDR q-value	0.275	0.099
Bandwidth	145.8	190.5
Controls	-	Y

Running Variable Control

Y

Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Each cell shows results for a separate regression. Each Panel shows results for a different dependent variable and the key independent variable is an indicator for policy enactment. Robust standard errors are in parentheses. All specifications control for the distance from policy enactment. The optimal (MSE) bandwidth is used to determine the sample for each separate regression. We compute the FDR-adjusted q-values using the method proposed by Anderson (2008), adjusting for our five different outcomes.

Table 4

Release Regression Discontinuity Results for Indigent and Non-Indigent Defendants

	Optimal Bandwidth	
	(1)	(2)
Panel A: Non-indigent Defendants		
Outcome: Non-financial Bond		
RD.Estimate	0.0237*	0.0220
	(0.0143)	(0.0143)
Observations	10474	9940
Outcome Mean	0.823	0.823
Bandwidth	142.6	135.2
Outcome: Pretrial Detention		
RD.Estimate	-0.0140	-0.0149
	(0.0105)	(0.0102)
Observations	10338	9896
Outcome Mean	0.0856	0.0863
Bandwidth	135.7	130.6
Panel B: Indigent Defendants		
Outcome: Non-financial Bond		
RD.Estimate	0.0603***	0.0695***
	(0.0228)	(0.0226)
Observations	7366	6168
Outcome Mean	0.460	0.465
Bandwidth	116.5	98.04
Outcome: Pretrial Detention		

RD_Estimate	-0.0505**	-0.0426**
	(0.0199)	(0.0173)
Observations	9708	9784
Outcome Mean	0.557	0.558
Bandwidth	126.0	127.6
Controls	-	Y
Running Variable Control	Y	Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Each cell represents results for a separate regression where the key independent variable is an indicator of policy enactment. Robust standard errors are in parentheses. All specifications control for the distance from policy enactment. The optimal (MSE) bandwidth is used to determine the sample for each separate regression. Panel A and B present results for non-indigent and indigent subgroups respectively.

Table 5

Pretrial Crime Regression Discontinuity Results for Indigent and Non-Indigent Defendants

	Optimal Bandwidth	
	(1)	(2)
Panel A: Non-indigent Defendants		
Outcome: Non-violent Pretrial Crime		
RD.Estimate	0.00663	0.00342
	(0.0106)	(0.0106)
Observations	14106	13534
Outcome Mean	0.120	0.119
Bandwidth	182.6	174.1
Outcome: Violent Pretrial Crime		
RD.Estimate	-0.00407	-0.00428
	(0.00425)	(0.00465)
Observations	15716	12300
Outcome Mean	0.0184	0.0181
Bandwidth	202.7	158.0
Outcome: Conviction		
RD.Estimate	-0.0252*	-0.0263*
	(0.0139)	(0.0143)
Observations	18452	16394
Outcome Mean	0.360	0.356
Bandwidth	236.0	210.7

Panel B: Indigent Defendants

Outcome: Non-violent Pretrial Crime

RD.Estimate	0.0174	0.0146
	(0.0116)	(0.0114)
Observations	11170	11170
Outcome Mean	0.102	0.102
Bandwidth	144.8	144.7
Outcome: Violent Pretrial Crime		
RD.Estimate	-0.000659	-0.000204
	(0.00362)	(0.00371)
Observations	16118	14632
Outcome Mean	0.0123	0.0122
Bandwidth	204.5	187.3
Outcome: Conviction		
RD.Estimate	-0.0129	-0.00834
	(0.0171)	(0.0167)
Observations	13072	12172
Outcome Mean	0.603	0.603
Bandwidth	168.7	157.4
Controls	-	Y
Running Variable Control	Y	Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Each cell shows results for a separate regression. Each Panel shows results for a different dependent variable and the key independent variable is an indicator for policy enactment. Robust standard errors are in parentheses. All specifications control for the distance from policy enactment. The optimal (MSE) bandwidth is used to determine the sample for each separate regression. Panel A and B present results for non-indigent and indigent subgroups respectively.

Table 6

Release Regression Discontinuity Results for White and Minority Defendants

	Optimal Bandwidth	
	(1)	(2)
Panel A: White Defendants		
Outcome: Non-financial Bond		
RD.Estimate	0.0351*	0.0286
	(0.0210)	(0.0178)
Observations	7514	7818
Outcome Mean	0.673	0.673
Bandwidth	121.8	125.2
Outcome: Pretrial Detention		
RD.Estimate	-0.0206	-0.0259*
	(0.0188)	(0.0138)
Observations	8364	8878
Outcome Mean	0.269	0.271
Bandwidth	125.6	133.2
Panel B: Minority Defendants		
Outcome: Non-financial Bond		
RD.Estimate	0.0541***	0.0526***
	(0.0206)	(0.0181)
Observations	7980	8240
Outcome Mean	0.641	0.641
Bandwidth	107.7	110.6
Outcome: Pretrial Detention		

RD_Estimate	-0.0392**	-0.0306**
	(0.0179)	(0.0144)
Observations	10924	10746
Outcome Mean	0.360	0.360
Bandwidth	126.7	124.6
Controls	-	Y
Running Variable Control	Y	Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Each cell represents results for a separate regression where the key independent variable is an indicator of policy enactment. Robust standard errors are in parentheses. All specifications control for the distance from policy enactment. The optimal (MSE) bandwidth is used to determine the sample for each separate regression. Panel A and B present results for white and minority defendants respectively.

Table 7

Pretrial Crime Regression Discontinuity Results for White and Minority Defendants

	Optimal Bandwidth	
	(1)	(2)
Panel A: White Defendants		
Outcome: Non-violent Pretrial Crime		
RD_Estimate	0.00114 (0.00875)	0.00195 (0.00856)
Observations	17538	18052
Outcome Mean	0.0866	0.0868
Bandwidth	257.3	265.3
Outcome: Violent Pretrial Crime		
RD_Estimate	-0.00397 (0.00461)	-0.00116 (0.00431)
Observations	9458	10390
Outcome Mean	0.0132	0.0133
Bandwidth	140.7	153.4
Outcome: Conviction		
RD_Estimate	-0.0368** (0.0186)	-0.0413*** (0.0140)
Observations	11282	17590
Outcome Mean	0.471	0.474
Bandwidth	167.2	258.3

Panel B: Minority Defendants

Outcome: Non-violent Pretrial Crime

RD.Estimate	0.0197*	0.0160
	(0.0106)	(0.0103)
Observations	14750	15026
Outcome Mean	0.122	0.122
Bandwidth	167.0	170.1
Outcome: Violent Pretrial Crime		
RD.Estimate	-0.00613	-0.00556
	(0.00473)	(0.00466)
Observations	12756	12526
Outcome Mean	0.0168	0.0168
Bandwidth	146.8	144.3
Outcome: Conviction		
RD.Estimate	-0.00609	-0.00937
	(0.0165)	(0.0145)
Observations	14146	16042
Outcome Mean	0.485	0.487
Bandwidth	161.9	181.3
Controls	-	Y
Running Variable Control	Y	Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Each cell shows results for a separate regression. Each Panel shows results for a different dependent variable and the key independent variable is an indicator for policy enactment. Robust standard errors are in parentheses. All specifications control for the distance from policy enactment. The optimal (MSE) bandwidth is used to determine the sample for each separate regression. Panel A and B present results for white and minority defendants respectively.

OHIO RISK ASSESSMENT SYSTEM: PRETRIAL ASSESSMENT TOOL (ORAS-PAT)				
Name: _____		Date of Assessment: _____		
Case #: _____		Name of Assessor: _____		
Pretrial Items				Verified
1.1	Age at First Arrest 0 = 33 or Older 1 = Under 33	<input type="text"/>	<input type="text"/>	<input type="text"/>
1.2	Number of Failure-to-Appear Warrants Past 24 Months 0 = None 1 = One Warrant for FTA 2 = Two or more FTA Warrants	<input type="text"/>	<input type="text"/>	<input type="text"/>
1.3	Three or more Prior Jail Incarcerations 0 = No 1 = Yes	<input type="text"/>	<input type="text"/>	<input type="text"/>
1.4	Employed at the Time of Arrest 0 = Yes, Full-time 1 = Yes, Part-time 2 = Not employed	<input type="text"/>	<input type="text"/>	<input type="text"/>
1.5	Residential Stability 0 = Lived at Current Residence Past Six Months 1 = Not Lived at Same Residence	<input type="text"/>	<input type="text"/>	<input type="text"/>
1.6	Illegal Drug Use during Past Six Months 0 = No 1 = Yes	<input type="text"/>	<input type="text"/>	<input type="text"/>
1.7	Severe Drug Use Problem 0 = No 1 = Yes	<input type="text"/>	<input type="text"/>	<input type="text"/>
Total Score:		<input type="text"/>		
Scores	Rating	% of Failures	% of Failure to Appear	% of New Arrest
0-2	Low	5%	5%	0%
3-5	Moderate	18%	12%	7%
6+	High	29%	15%	17%

Figure 1

Ohio Risk Assessment Score in Travis County

Notes: This figure shows the risk assessment tool used in Travis County, Texas.

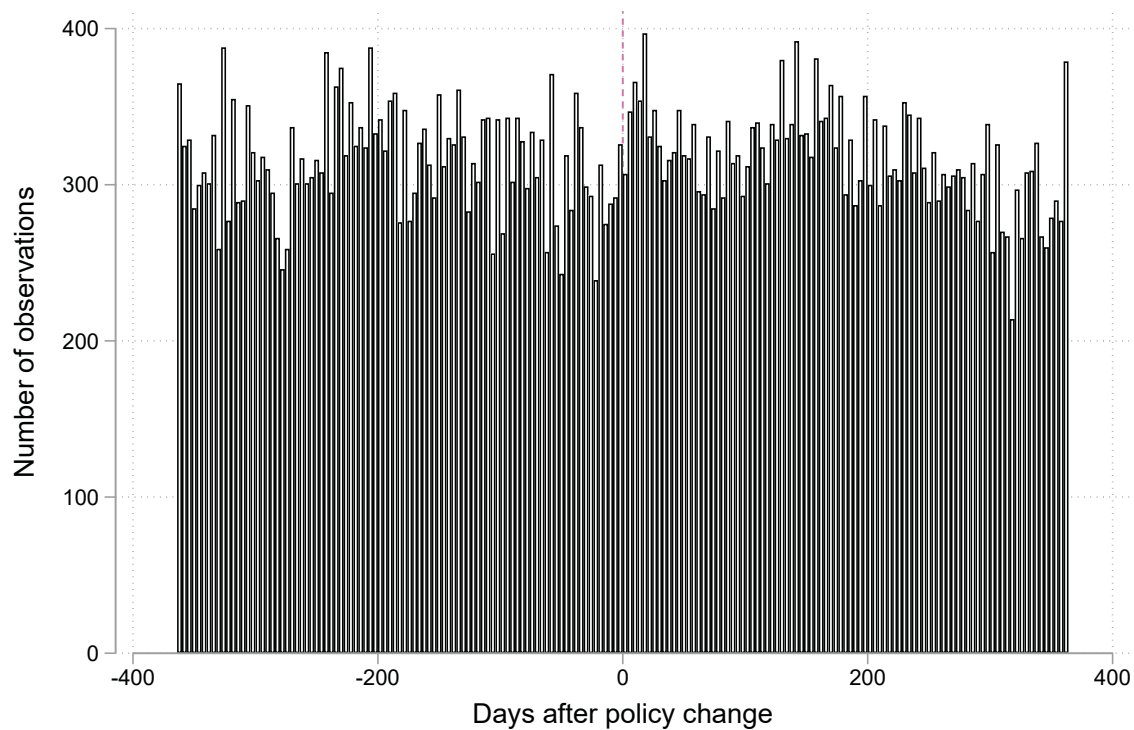
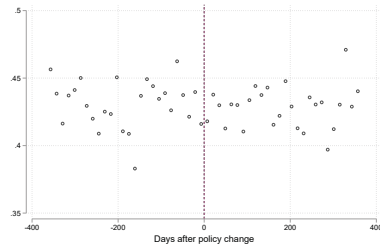


Figure 2

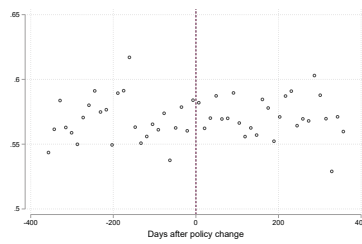
Frequency of Cases Across Time

Notes: This figure shows the distribution of running variable observations near the adoption of risk assessment scores. Each bin is 2 days. The dashed line marks the day of the policy change.

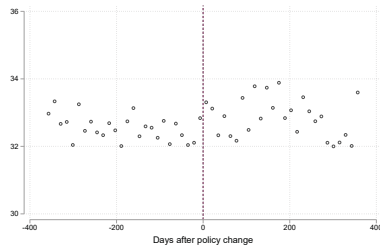
Panel A: White Defendant



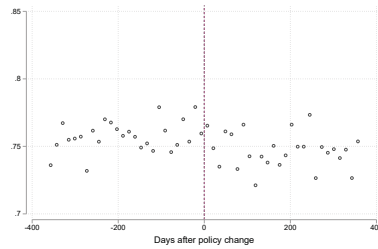
Panel B: Minority Defendant



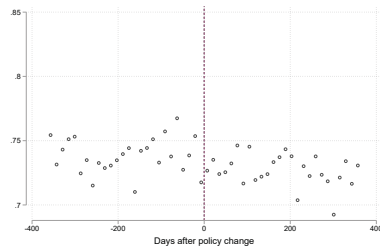
Panel C: Defendant Age



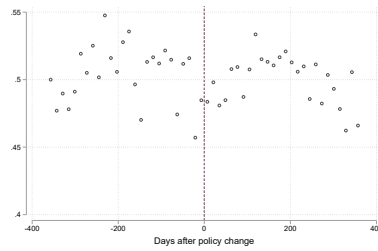
Panel D: Defendant Gender



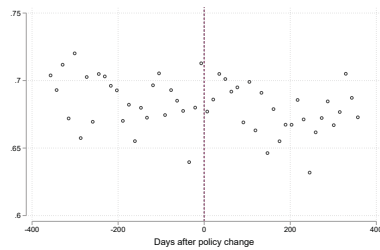
Panel E: No Priors



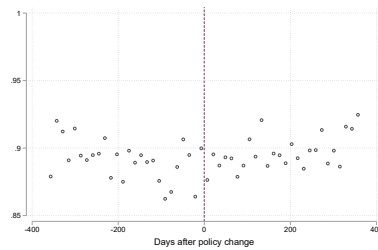
Panel F: Indigent Status



Panel G: Misdemeanor or Felony Case



Panel H: United States Citizen



Panel I: Mental Health Flag

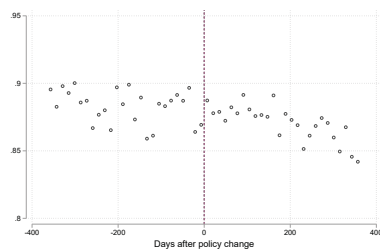
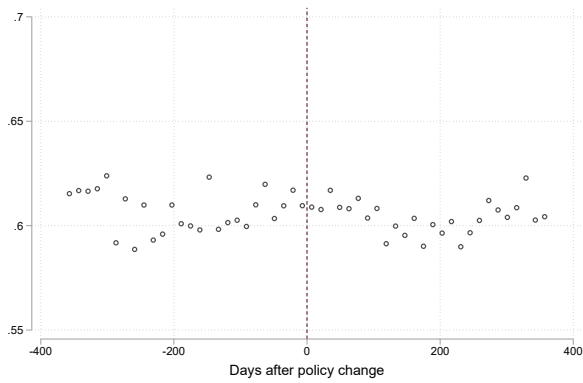


Figure 3

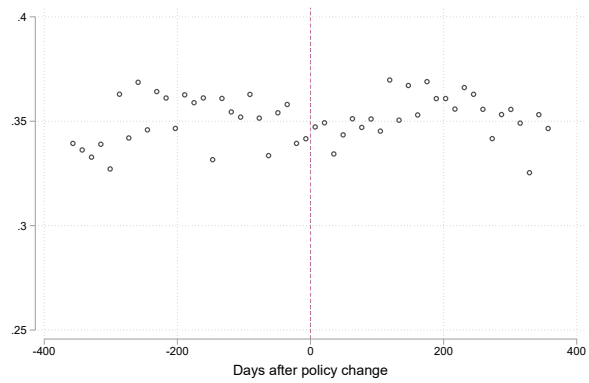
Smoothness of Baseline Covariates

Notes: These figures plot tests of the regression discontinuity design. Each figure plots means of the outcome variable in two week bins.

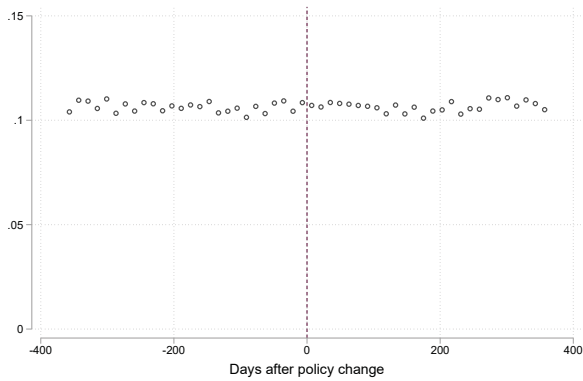
Panel A: Predicted Probability of Release on Non-financial Bond



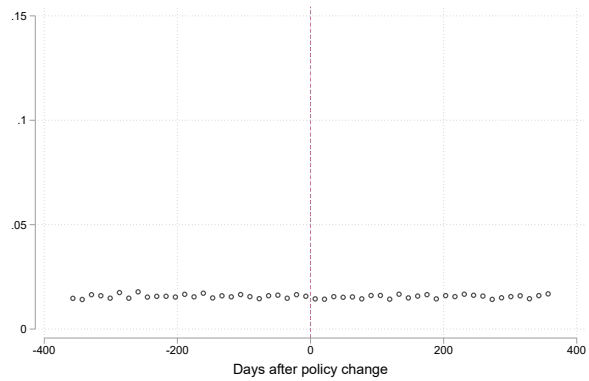
Panel B: Predicted Probability of Pretrial Detention



Panel C: Predicted Probability of Non-Violent Pretrial Crime



Panel D: Predicted Probability of Violent Pretrial Crime



Panel E: Predicted Probability of Conviction

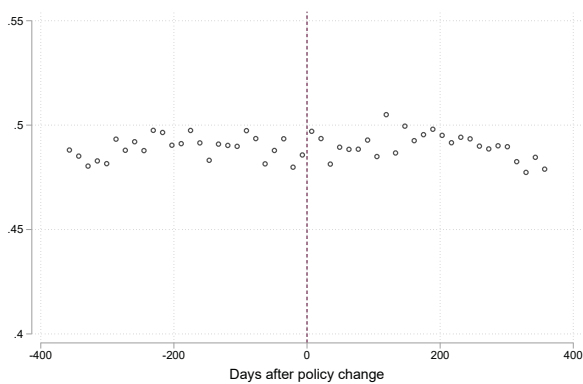


Figure 4

Regression Discontinuity Results for Predicted Values

Notes: These figures plot tests of the regression discontinuity design. Each figure plots means of the outcome variable in two week bins. Outcome variables are predicted using observable case and defendant characteristics. Specifically, we use race, age, gender, criminal history, indigent status, severity of arrest, mental health status, and US citizenship status, along with a court and day-of-week fixed effects. A bandwidth of 360 days is shown.

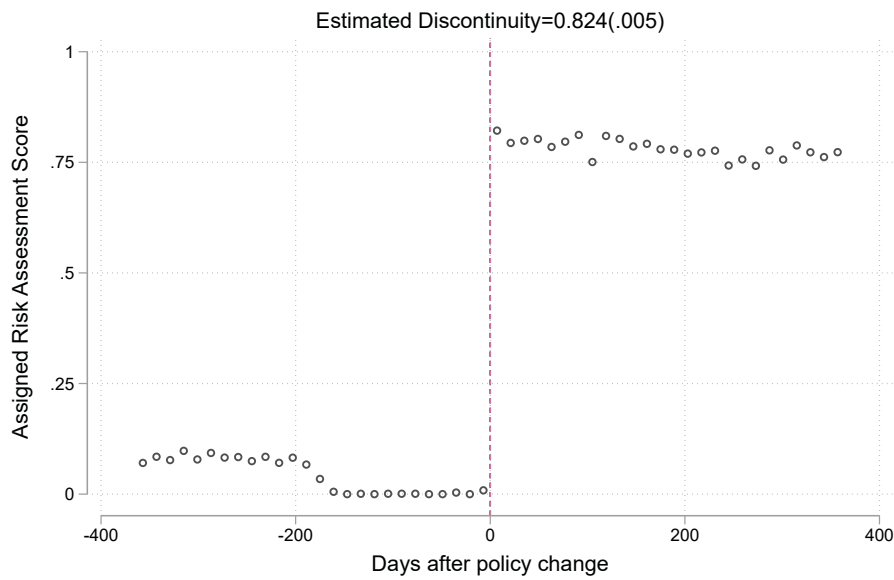
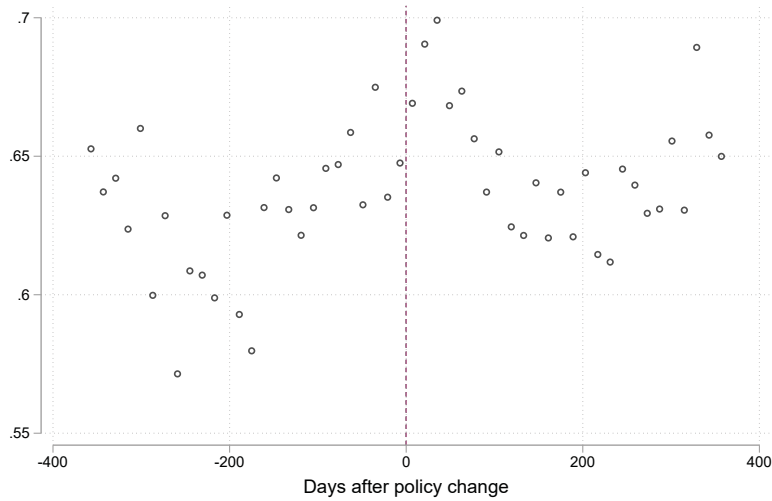


Figure 5

Regression Discontinuity Results for the Probability of Receiving a Risk Assessment Score

Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on score usage by plotting the mean of risk assessment score take-up in two week bins. The outcome variable takes on a value of one if a defendant has a risk assessment score and zero if she does not. A bandwidth of 360 days is shown. There was a small pilot study run about a year before the policy change.

Panel A: Non-financial Bond



Panel B: Pretrial Detention

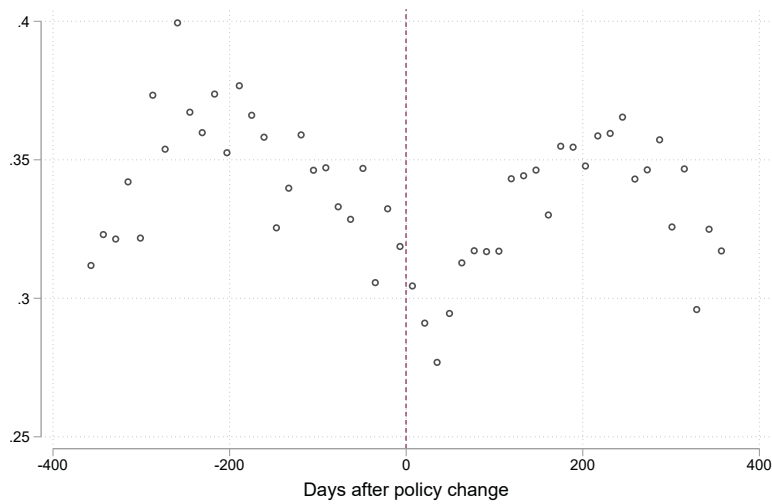
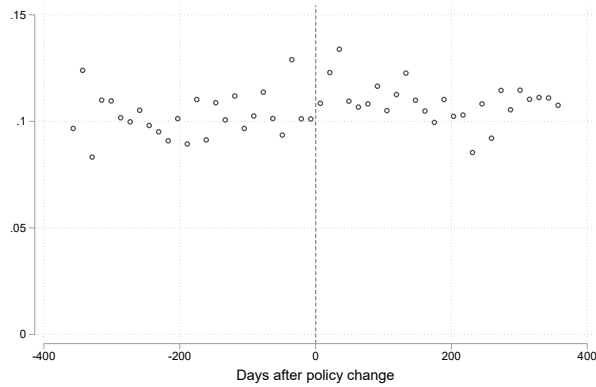


Figure 6

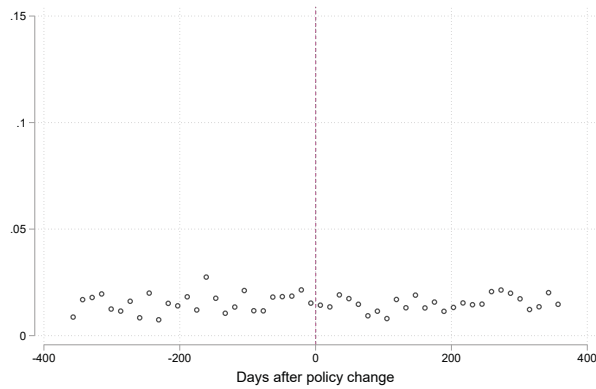
Regression Discontinuity Results for Non-financial Bond and Pretrial Detention

Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on non-financial bond or pretrial detention by plotting the mean non-financial bond or pretrial detention in two week bins. A bandwidth of 360 days is shown.

Panel A: Probability of Non-Violent Pretrial Crime



Panel B: Probability of Violent Pretrial Crime



Panel C: Probability of Conviction

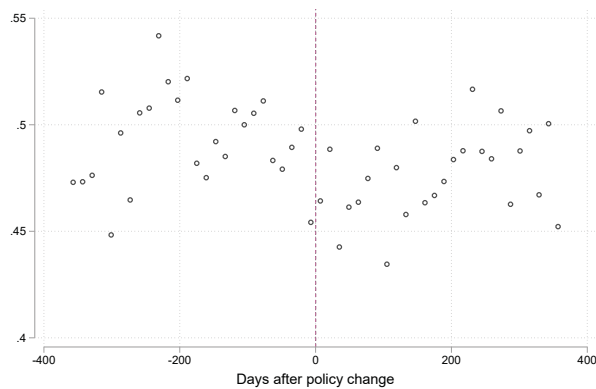
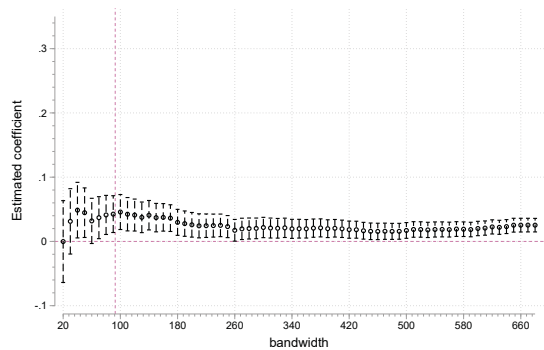


Figure 7

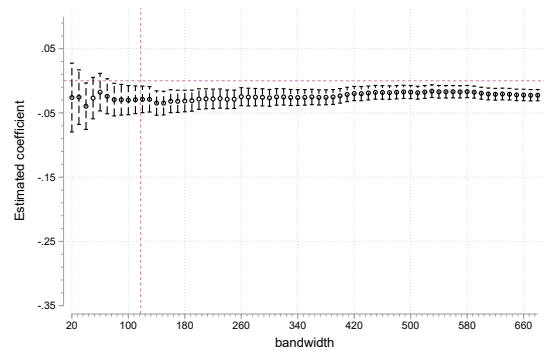
Regression Discontinuity Results for Conviction and Pretrial Crime

Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on conviction, non-violent and violent pretrial crime by plotting the mean of the outcome variable in two week bins. A bandwidth of 360 days is shown.

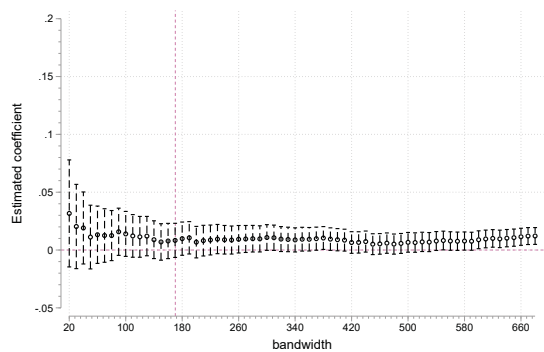
Panel A: Linear - Non-financial Bond



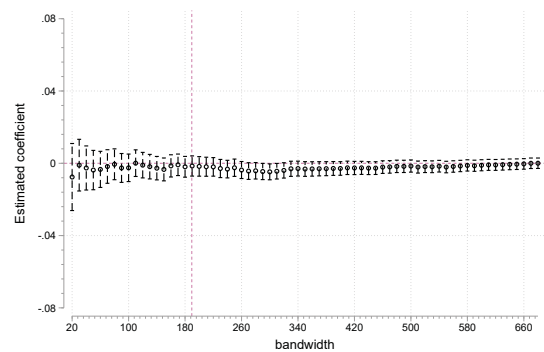
Panel B: Linear - Pretrial Detention



Panel C: Linear - Non-violent Pretrial Crime



Panel D: Linear - Violent Pretrial Crime



Panel E: Linear - Conviction

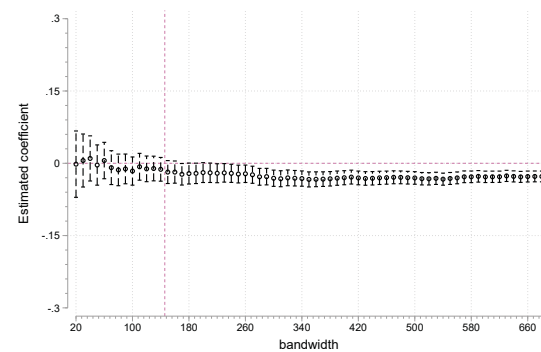
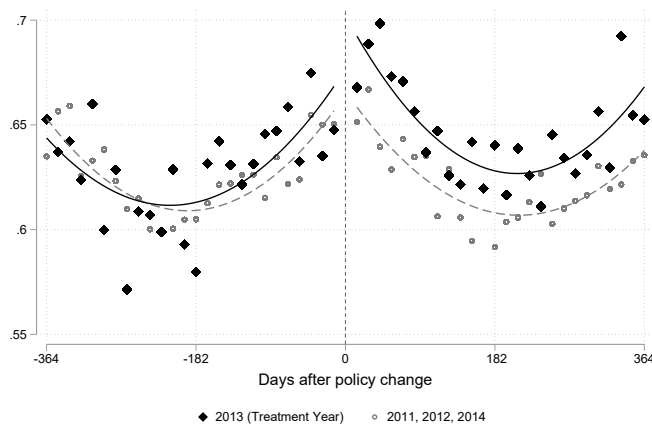


Figure 8

Non-financial and Pretrial Detention Bandwidth Robustness

Notes: Each figure plots coefficients from 64 different regressions using different bandwidths. Ninety-five percent confidence intervals are also presented. The optimal MSE bandwidth for each specification is marked with the dashed line.

Panel A: Non-financial Release



Panel B: Pretrial

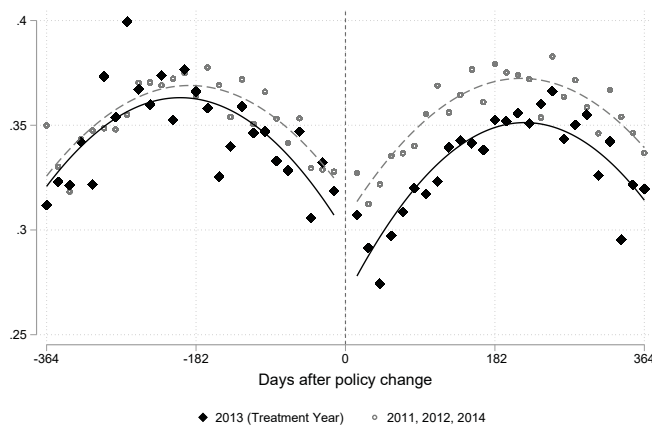
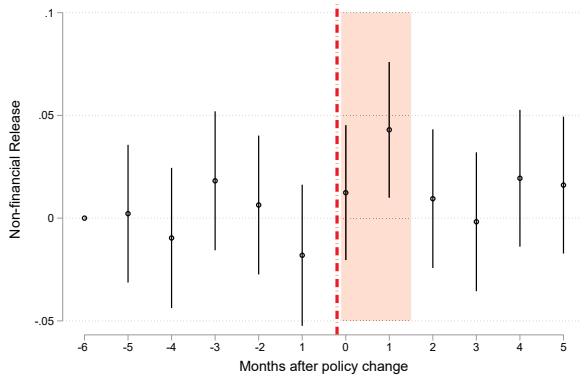


Figure 9

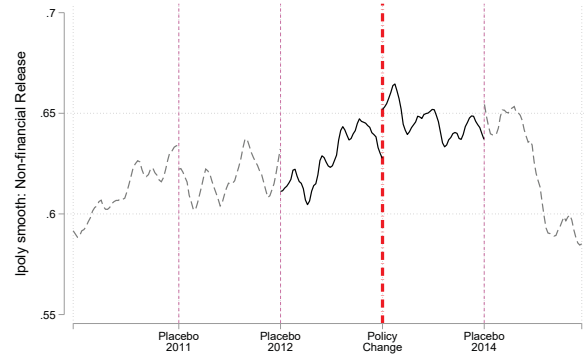
Robustness to Seasonality - Regression Discontinuity Results for Conviction and Pretrial Crime

Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on non-financial bond or pretrial detention by plotting the mean non-financial bond or pretrial detention in two week bins. A bandwidth of 360 days is shown. The gray dots also show the mean of the outcome for 2011, 2012, and 2014 (years with no treatment).

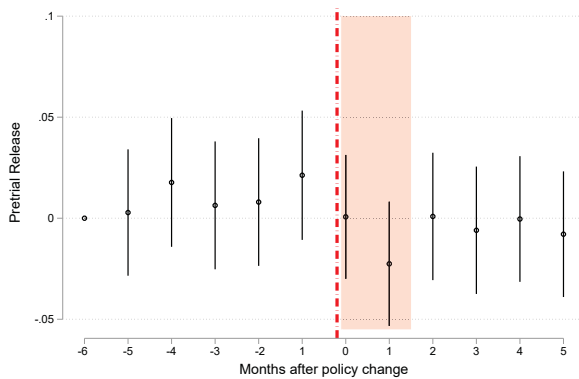
Panel A: Non-Financial Release



Panel B: Non-Financial Release



Panel C: Pretrial Detention



Panel D: Pretrial Detention

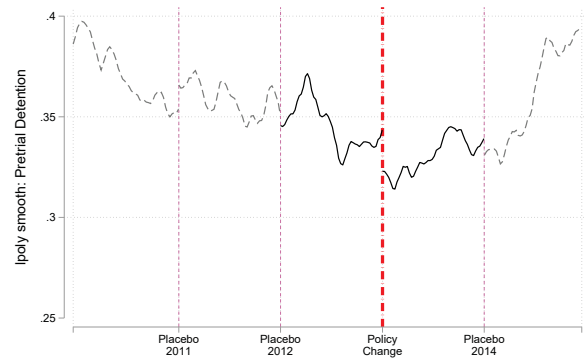


Figure 10

Dynamic Effects of Risk Assessment Scores

Notes: Figures 10 (a) and (c) show coefficients from the regression of non-financial bond, and pretrial detention on indicators for months before or after risk assessment adoption interacted with an indicator for our true treatment time period. We also include fixed effects for months before/after “treatment” and each time period. Individual level controls for race, age, gender, citizenship, and indigent status of the defendant along with controls for the severity of the crime (misdemeanor or not) as well as fixed effects for the court assigned and day-of-week of booking are used. Figures (b), and (d) show the time trend for each outcome estimated using local linear smoothing with a bandwidth of 14 days controlling for seasonality using week of the year fixed effects (i.e., using STATA’s *lpoly*). The x-axis indicates the booking date. Pink dashed lines mark dates of the policy change in other years. The red dashed line marks the true policy change.