

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

CarlyWill Sloan, George Naufal, and Heather Caspers *

December 29, 2019

Abstract

The use of risk assessment scores as a means of decreasing pretrial detention for low-risk, primarily poor defendants is increasing rapidly across the United States. Despite this, there is little evidence on how risk assessment scores alter criminal outcomes. Using administrative data from a large county in Texas, we estimate the effect of a risk assessment score policy on judge bond decisions, defendant pretrial detention, and pretrial crime. We identify short-term effects by exploiting a large, sudden policy change using a regression discontinuity design. This approach effectively compares defendants booked just before and after the policy change. Results show that adopting a risk assessment score leads to increased release on non-financial bond and decreased pretrial detention. These results appear to be driven by poor defendants, do not worsen existing racial disparities, and deteriorate with time. Additionally, we find risk assessment scores did not increase violent pretrial crime, however, there is some suggestive evidence of small increases in non-violent pretrial crime.

*Sloan: Texas A&M University, cwsloan_1992@tamu.edu; Naufal: Texas A&M University and IZA, gnafaul@tamu.edu; Caspers: Texas A&M University, hcaspers@ppri.tamu.edu

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.

1 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 the poor and those with low risk, many jurisdictions are beginning to look for ways to reduce their pretrial population. A common suggestion is a shift 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. Assessing defendant risk is not a new idea in criminal justice, but in recent years risk assessment has taken on the additional meaning of using more technical and actuarial methods of predicting the likelihood of future crimes or failure to appear. Supporters of risk assessment scores argue that assessing individuals based on their risk rather than income could lead to less pretrial detention, allowing defendants to keep their jobs and imposing near zero costs on the criminal justice system if defendants do not commit new crimes pretrial. These policies are also most likely to benefit low-income defendants who cannot post bail. Opponents claim that increasing pretrial release through the use of non-financial bond (release on your own recognizance) and risk assessment scores could increase pretrial crime as the expected penalty of future crime is lower, further threatening societal safety and raising costs. It is also possible that risk assessment scores could exacerbate existing racial disparities, as these scores often include components, i.e. employment or criminal history,

that are correlated with defendant's race.¹ This paper focuses on the question of whether the use of risk assessment scores can increase non-financial bond, and decrease pretrial detention without increasing societal costs.

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 scores. 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 effect of risk assessment 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 booked after arrest. This type of sudden change can be used to identify local effects through a regression discontinuity design. Using the timing of the policy change, we are able to compare defendants booked just

¹Academics and journalists express varying degrees of concern about potential racial biases in risk assessment scores (e.g. Angwin et al., 2019; Doleac et al., 2017).

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.

First we present results for release outcomes. We show that the use of a risk assessment score increases release on non-financial bail by 4.5-7% and decreases pretrial detention by 7-9.5%. We also provide evidence that results are driven by low-income defendants and do not worsen existing racial disparities. Second, we are able to rule out meaningful increases in violent pretrial crime. 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. It also appears judges returned to their old release patterns about three months after the policy change.

To our knowledge, this paper provides some of the first causal evidence on the effects of pretrial risk assessment scores. 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 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 et al., 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

et al., 2009; Zhang et al., 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. This paper is most similar in spirit to Stevenson (2018b). She evaluates multiple pretrial risk assessment score policy changes in Kentucky using an event study framework. Stevenson (2018b) 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 pretrial crime. Further, using data from a large metropolitan area allows us to consider whether risk assessment scores altered existing racial disparities in pretrial detention and non-financial bonds in a more diverse area and specifically for low-income 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 et al., 2018; Heaton et al., 2017; Leslie and Pope, 2016; Stevenson, 2018a). Others have considered the effect of nonmonetary bail on outcomes, finding that nonmonetary bail decreases conviction rates (Gupta et al., 2016).

The results of this paper have important implications for criminal justice actors and defendants. First, our finding that risk assessment scores increase non-financial bail and

decrease pretrial detention suggests that this policy can be used to lower costs at least in the short term. These savings could be substantial as the estimated annual cost of pretrial detention in the US is \$13.4 billion (Wagner and Rabuy, 2017). Significantly, we also show that this reduction in pretrial detention and increase in non-financial bail releases could be possible without increases in violent pretrial crime.

Second, the use of risk assessment scores is important to defendants because it relieves low-income and minority defendants of the potentially disproportionate burden of financial bail and therefore pretrial detention. Perhaps most importantly, decreases in pretrial detention are also associated with greater job stability, less reliance on government assistance, lower probability of conviction, and less separation from family (Dobbie et al., 2018; Stevenson, 2018a). To the extent that our results apply in other settings, these findings indicate that risk assessment score policies may be an effective tool for decreasing the income-based disparity in pretrial detention and improving the lives of defendants at least in the short term. Notably, these decreases in pretrial detention are not associated with increases in violent pretrial crime, implying minimal risk and costs to society. However, policy makers must be careful to weigh these potential benefits with the possibility for some increases in non-violent pretrial crime. They should also recognize the benefits of risk assessment scores may be short-lived and are not likely to completely overhaul an existing bail system.

2 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.

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 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.

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 bail or if they pay for release², 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 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.

²A defendant can either pay their bail in full or a bail bondsman can post bail instead.

3 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.³ Unfortunately, information on non-financial bonds is missing for roughly 11 percent (15,188) 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 5.6.

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.⁴ Conviction is recorded as a one if a defendant is convicted of the crime

³It 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.

⁴Our results are similar in significance and magnitude if we define pretrial detention as being in jail for more than one or three days.

they were originally arrested for, and zero otherwise. Pretrial crime is measured for all defendants, regardless of their pretrial bond or detention status. 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 percent, 76, and 89 percent respectively.⁵ 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.⁶ For the entire time period, 33 percent of defendants have a risk assessment score recorded, although 76 percent of defendants have a risk assessment score after January 2013.

4 Methods

4.1 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,

⁵Travis 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.

⁶This is the definition of indigence from Travis County Criminal Courts (2012).

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.

Formally, we estimate the following individual-level OLS model:

$$Outcome_{it} = \beta_0 + \beta_1 policyenacted_t + f(daysfromcutoff)_t + \lambda_i + \pi_c + \gamma_d + \epsilon_{it}$$

Here, i indexes individual defendants and t the date of booking. $Policyenacted_t$ 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, $daysfromcutoff_t$, is defined as days from the date of policy enactment, or $dateofbooking_t - policyenactmentdate_t$, and the function $f(\cdot)$ captures the underlying relationship between the outcome variable and the running variable. By interacting $f(\cdot)$ with $policyenacted_t$, we allow the slopes of our fitted lines to differ on either

side of the policy change. Our coefficient of interest, β_1 , captures the intent-to-treat effect of the risk assessment score policy. λ_i contains individual-level controls that could alter the precision of our estimates, but should not drastically change our estimates of β_1 if our identifying assumption holds. π_c is court-specific fixed effects, and γ_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. In our results, we also control for the running variable, $daysfromcutoff_t$, in many ways by allowing $f(\cdot)$ to take on different forms. Our preferred specification defines $f(\cdot)$ as a linear function because it enforces the least functional form assumptions on the data. Finally, we report robust standard errors calculated as suggested in Calonico et al. (2017).⁷

4.2 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

⁷Although 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 remains the same for 5 of the 6 estimates presented in Table 2. One estimate is significant at the 5% level instead of the 1% due to clustering. For pretrial detention, all six of the estimates in Table 2 retain their significance level with clustering.

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 Appendix Table 1 show the results for this test. There is only one small visible jump in defendant and case characteristics in the graphs presented (Defendant Age). Of the 20 estimates presented in Appendix Table 1, 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 A2 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.

5 Results

5.1 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 30-day bins and a linear fit of the outcome, which is allowed to vary on each side of the policy change threshold. 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.⁸ 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. 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.

⁸Risk assessment usage was greater than zero for a few months before January 2013 because Travis County elected to run a pilot study.

5.2 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 30-day bins and a linear fit of the outcome, which is allowed to vary on each side of the policy change threshold. This figure provides visual evidence that implementing risk assessment scores increases the likelihood of release on non-financial bond and decreases pretrial detention. 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. Each column includes 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).⁹ As with every specification of a regression discontinuity, the polynomials are allowed to differ on each side of the policy change cutoff. Even numbered columns allow the running variable, days to the cutoff, to vary quadratically and odd numbered columns present linear results. Columns (1)-(2) present results for double the MSE optimal bandwidth, columns (3)-(4) 1.5 times the MSE optimal bandwidth, and columns (5)-(6) the optimal bandwidth. If our identifying assumption holds, we would expect that our coefficient of interest would remain similar in magnitude. Across all eight columns, our estimates remain statistically significant at conventional levels and are of similar magnitudes for non-financial bond and pretrial detention.¹⁰

Our estimates for non-financial bond range from 0.029 to 0.043. These results indicate that the implementation of a risk assessment score policy increases the likelihood of release on non-financial bond by about 3-5 percentage points (4.5-7%). For pretrial detention, our estimates range from -0.025 to -0.033, showing the risk assessment score policy decreases the chance of pretrial detention by about 3 percentage points (7-9.5%).

Because our results include four 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 least the five percent level for all but

⁹The primary specification in this paper includes controls. If we recreate our results with no controls our estimates are not meaningfully different. For example our results for non-financial bond range from 0.0271 to 0.0428. Pretrial detention estimates range from -0.0172 to -0.038

¹⁰Our results are similar in significance and magnitude if we define pretrial detention as being in jail for more than one or three days.

one of the specifications. 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 detainment 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 and functional form 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 A1 Panels (a) and (d) show the coefficients and standard errors from each model for non-financial bond and pretrial detention. We also complete the same exercise, but with allowing the running variable to vary quadratically and cubically. The results for these models are shown in Figure A1 Panels (b), (c), (e), and (f). 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 and functional form.

We also conduct a permutation test in the spirit of Abadie et al. (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 A2. Nearly all placebo coefficients (97.99) 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 95.88 percent of our placebo estimates. These simulations imply p-values of 0.02 and 0.04 for our non-financial bond and pretrial detention estimates, indicating that our results cannot be explained by serial correlation. Further, these simulations imply our estimates are not simply due to chance.

Finally, it is natural to wonder what type of defendants were released after the adoption of a risk assessment score policy. To investigate this question, we estimate the effect of a risk assessment score on multiple subgroups in Figure 7. Each figure shows the regression discontinuity estimate from a separate regression. Estimates for the optimal bandwidth are reported.

First we show results for defendants with low predicted risk versus high predicted risk. Because defendants booked before the adoption of the risk assessment score are not assigned a score, we use all case observables (race, age, gender, criminal history, indigent status, severity of arrest, mental health status, US citizenship status, and specific court) to predict a risk score for every defendant. Defendants with above average predicted risk are considered high risk and those with below average predicted risk are considered low risk. Overall it appears that defendants with higher predicted risk were more likely to be offered non-financial bond

or not detained pretrial than were low risk defendants. Older, felony-committing, minority, and indigent defendants are all more likely to be released on non-financial bond compared to younger, misdemeanor-committing, white, and non-indigent defendants.

Together these results show the adoption of a risk assessment scores policy caused increased offers of non-financial bonds by judges, which appears to lead to meaningful decreases in pretrial detention. Since we find results for pretrial detention, we also consider whether a risk assessment score policy alters pretrial crime and conviction.

5.3 Effects of a Risk Assessment Score Policy on Pretrial Crime

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 8. All graphs in Figure 8 show the mean of the outcome variable in 30-day bins and a linear fit of the outcome, which is allowed to vary on each side of the policy change threshold. Figure 8(a) and Figure 8(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 allow the running variable to vary quadratically and odd columns are linear. Each column controls for defendant race, age, gender, citizenship, mental health status, and indigent status, as well as the severity of the crime (misdemeanor or not). 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 across all six columns. For non-violent pretrial crime, estimates range from 0.0095 to 0.01 across the

table. Only three estimates are significant at the ten percent level. Although there appears to be some evidence of small increases in non-violent pretrial crime, our results are not robust to alternative specifications.

Next we consider violent pretrial crime . Across all columns our estimates remain stable, ranging from -0.001 to -0.005. We are also able to rule out any increase in violent pretrial crime when using the larger sample size from twice the optimal bandwidth.¹¹ We also report FDR q-values for pretrial crime outcomes.¹² 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 and functional forms. 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 A3 shows the coefficients and standard errors from each model for non-violent and violent pretrial crime. We also complete the same exercise, but with allowing the running variable to vary quadratically and cubically. The results for these models are shown in Figure A3 Panels (b), (c), (e), and (f). The dashed lines represent the optimal MSE bandwidth. The estimated coefficients remain consistent across the different bandwidths. 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 et al.,

¹¹-.000748 is the top of the 95% confidence interval from this specification.

¹²Again, we correct for 5 categories.

2018; Stevenson, 2018a). Figure 8(c) shows our results for conviction. Here there appears to be some decrease in conviction after the adoption of a risk assessment score. Corresponding estimates are found in Table 3. All six coefficients are negative and suggest that risk assessment scores could have decreased the odds of conviction for defendants. However, the magnitude and significance of the coefficients are not robust to multiple specifications.¹³ Therefore, we do not consider this outcome for subgroup analysis.¹⁴

5.4 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 9. 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 A3.

In Table A3, Panel A presents results for indigent defendants and Panel B shows re-

¹³Further, Figures A3(g), A3(h) and A3(i) show conviction results for different bandwidths.

¹⁴For all subgroups (indigent, non-indigent, minority, white) our estimates are not robust to different specifications and bandwidths.

sults for non-indigent defendants. Similar to earlier result tables, each specification includes all case controls. Even columns allow the running variable to vary quadratically and odd columns are linear. Across each specification the coefficient for non-financial bond and pre-trial 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. 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 10. Panels (a) and (d) repeat the entire sample results for comparison. Indigent results are shown in Panels (b) and (e), while Panels (c) and (f) 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.

Table A4 shows pretrial crime 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. Further, for each subgroup the coefficient for violent pretrial crime is negative, again suggesting there are no increases in violent pretrial crime across either group.

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.

5.5 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.¹⁵ 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 11. 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 A5. For non-financial bail, in each specification, the magnitude of our estimate is larger for minority defendants than white defendants.

We also consider pretrial crime for minority and white defendants in Figure 12. Results for non-violent and violent pretrial crime are shown in Figure 12. Panels (a) and (d) repeat

¹⁵In this setting minority defendants are non-white and Hispanic. White defendants are only white.

the entire sample results for comparison. White defendant results are show in Panels (b) and (e), while Panels (c) and (f) 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 however, there appears to be no increase in pretrial crime for white or minority defendants.

Table A6 shows pretrial crime 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. Further, for each subgroup the coefficient for violent pretrial crime is negative, again suggesting there are no increases in violent pretrial crime across either group.

Taken together these results show that our results for non-financial bail 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 bail or pretrial detention. If anything, the magnitude of our coefficients suggests that risk assessment score may have decreased racial disparities.

5.6 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 A4. 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 A7. Even columns allow the running variable to vary quadratically and odd columns are linear. Columns (1)-(6) use the optimal bandwidth determined in Table 2 for release on non-financial bond. Columns (7)-(8) use the optimal bandwidth for the probability of missing data. Across all eight columns, the coefficient remains statistically insignificant and close to zero. 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 A5 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.

5.7 Long Term Effects

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 conduct event study analysis with results presented in Figure A6.¹⁶ Visually, it is clear that the effect of non-financial bonds lasts for only the first two months after the policy change and that the rate of release on non-financial bonds returns to non-financial bond just afterward. It is natural to wonder why we see such a short-lived effect from the risk assessment scores.

First, we note that Stevenson (2018b) also provided some evidence that the effects of risk assessment scores fade with time, so this is not an uncommon pattern. Travis County Pretrial Services also noted that 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.

¹⁶Formally we regress probability of release on non-financial bond on indicators months before and after the policy change in two-month bins. Our regression also controls for race, age, gender, citizenship, mental health status, and indigent status of the defendant, along with controls for the severity of the crime (misdemeanor or not) and fixed effects for the assigned court and booking day of the week. Finally, we add a court-specific time trend.

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.

6 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 risk assessment scores 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 4.5-7% increase in non-financial bonds and a 7-9.5% 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, it is possible that effects are only short-lived. Even with this qualification, 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.

References

- Abadie, A., A. Diamond, and J. 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, L., M. McManus, D. Brian, and D. P. 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, M. 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, J., J. Larson, L. Kirchner, and S. Mattu (2019). Machine bias. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Bureau of Justice Statistics (2013, Dec). Felony defendants in large urban counties 2009 statistical tables. *Bureau of Justice Statistics (BJS)*.
- Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). rdrobust: Software for regression discontinuity designs. *Stata Journal* 17(2), 372–404.
- Carmichael, D., G. Naufal, S. Wood, H. Caspers, and M. Marchbanks (2017). Liberty and justice: Pretrial practices in Texas.

Chanenson, S. L. and J. M. Hyatt (2016). The use of risk assessment at sentencing: Implications for research and policy. *Hyatt, JM & Chanenson, SL (2016). The Use of Risk Assessment at Sentencing: Implications for Research and Policy. Bureau of Justice Assistance, Washington, DC.*

Craver, J. (2017, Mar). Travis county: No place for bondsmen. *Austin Monitor*.

DeMichele, M., P. Baumgartner, M. Wenger, K. Barrick, M. Comfort, and S. Misra (2018). The public safety assessment: A re-validation and assessment of predictive utility and differential prediction by race and gender in Kentucky.

Didwania, S. H. (2018). The immediate consequences of pretrial detention: Evidence from federal criminal cases. *Working Paper*.

Dobbie, W., J. Goldin, and C. 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, J. L., M. Stevenson, and J. L. Doleac (2017, Aug). Are criminal risk assessment scores racist? <https://www.brookings.edu/blog/up-front/2016/08/22/are-criminal-risk-assessment-scores-racist/#7s8d6f87>.

Dressel, J. and H. Farid (2018). The accuracy, fairness, and limits of predicting recidivism. *Science advances* 4(1), eaao5580.

Flores, A. W., A. M. Holsinger, C. T. Lowenkamp, and T. 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, W. M., D. H. Zald, B. S. Lebow, B. E. Snitz, and C. Nelson (2000). Clinical versus mechanical prediction: a meta-analysis. *Psychological assessment* 12(1), 19.

Gupta, A., C. Hansman, and E. Frenchman (2016). The heavy costs of high bail: Evidence from judge randomization. *The Journal of Legal Studies* 45(2), 471–505.

Heaton, P., S. Mayson, and M. Stevenson (2017). The downstream consequences of misde-meanor pretrial detention. *Stan. L. Rev.* 69, 711.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2017). Human decisions and machine predictions. *The Quarterly Journal of Economics* 133(1).

Latessa, E. J., R. Lemke, M. Makarios, and P. Smith (2010). The creation and validation of the Ohio risk assessment system (ORAS). *Fed. Probation* 74, 16.

Leslie, E. and N. G. Pope (2016). The unintended impact of pretrial detention on case outcomes: Evidence from NYC arraignments.

Meredith, T., J. C. Speir, and S. 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.
<https://njcourts.gov/courts/assets/criminal/2018cjannual.pdf?c=taP>.

Schmidt, N., E. Lien, M. Vaughan, and M. 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, J. (2012, April). Keeping people in jail costs the county money, but is it in the best interest of public safety? *The Austin Chronicle*.

Stevenson, M. (2018a). Distortion of justice: How the inability to pay bail affects case outcomes. *Journal of Law, Economics and Organization, Forthcoming*.

Stevenson, M. T. (2018b). Assessing risk assessment in action. *Minnesota Law Review Forthcoming*.

Travis County Criminal Courts (2012). Travis County criminal courts Fair Defense Act program.

Turner, S., J. Hess, and J. 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. *United States Census Bureau*.

United States Census Bureau (2018, July). County population totals and components of change: 2010-2017. <https://www.census.gov/data/datasets/2017/demo/popest/counties-total.html>.

Wagner, P. and B. Rabuy (2017, Jan). Following the money of mass incarceration.

Zhang, S. X., R. E. Roberts, and D. Farabee (2014). An analysis of prisoner reentry and parole risk using compas and traditional criminal history measures. *Crime & Delinquency 60*(2), 167–192.

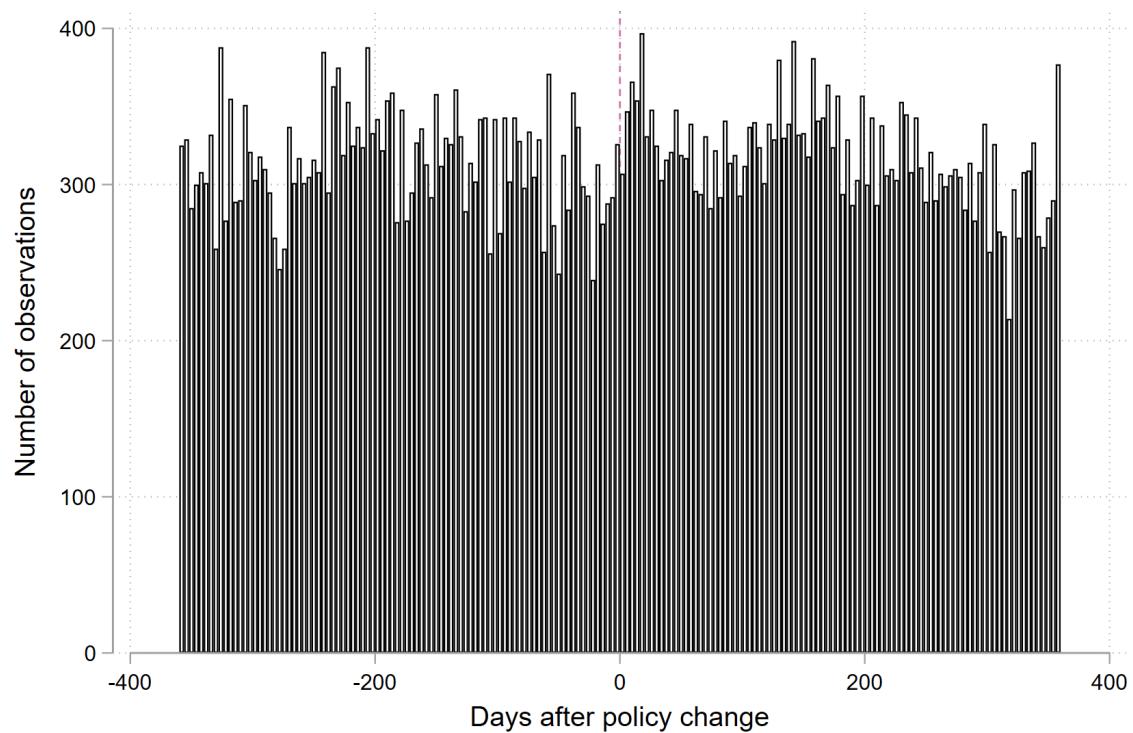
Figures and Tables

Figure 1: Ohio Risk Assessment Score in Travis County

OHIO RISK ASSESSMENT SYSTEM: PRETRIAL ASSESSMENT TOOL (ORAS-PAT)				
Name:	Date of Assessment: _____			
Case #:	Name of Assessor: _____			
Pretrial Items				
1.1 Age at First Arrest 0 = 33 or Older 1 = Under 33	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Verified	
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="checkbox"/>	<input checked="" type="checkbox"/>		
1.3 Three or more Prior Jail Incarcerations 0 = No 1 = Yes	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
1.4 Employed at the Time of Arrest 0 = Yes, Full-time 1 = Yes, Part-time 2 = Not employed	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
1.5 Residential Stability 0 = Lived at Current Residence Past Six Months 1 = Not Lived at Same Residence	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
1.6 Illegal Drug Use during Past Six Months 0 = No 1 = Yes	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
1.7 Severe Drug Use Problem 0 = No 1 = Yes	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
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%

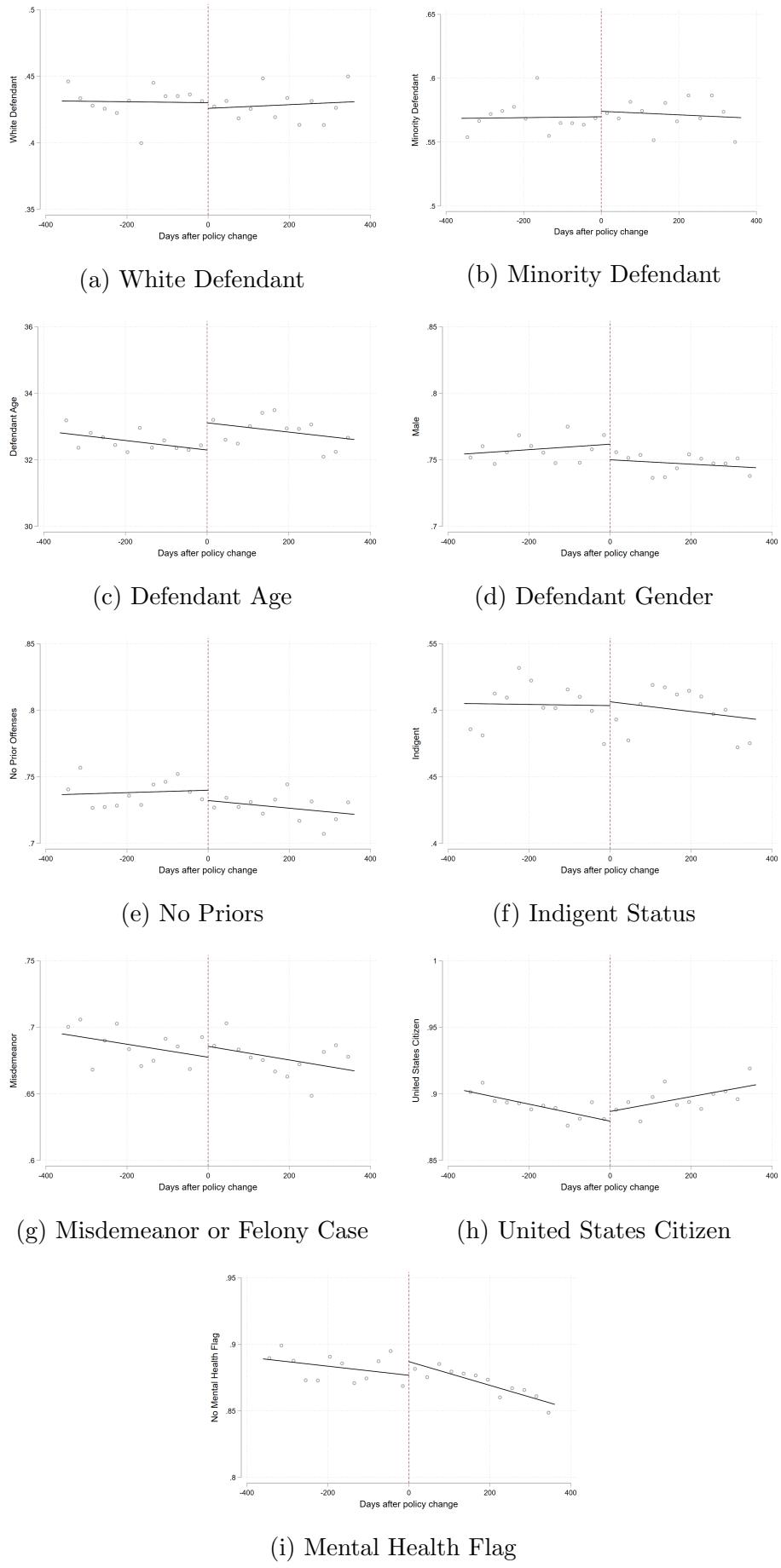
Notes: This figure shows the risk assessment tool used in Travis County, Texas.

Figure 2: Frequency of Running Variable



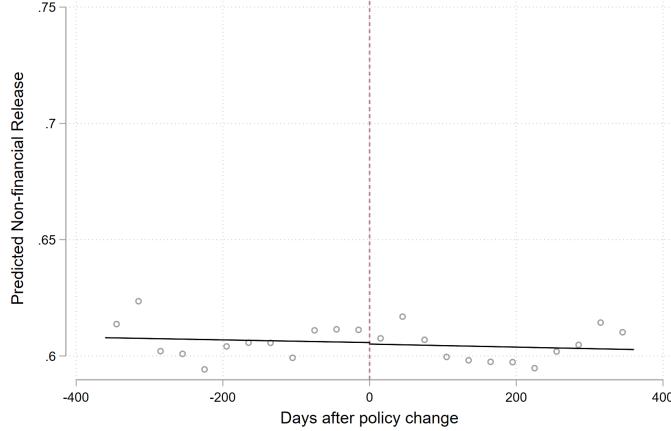
Notes: This figure shows the distribution of running variable observations near the adoption of risk assessment scores. Each bin is 2 days. The dashed line marks the day of the policy change.

Figure 3: Smoothness of Baseline Covariates

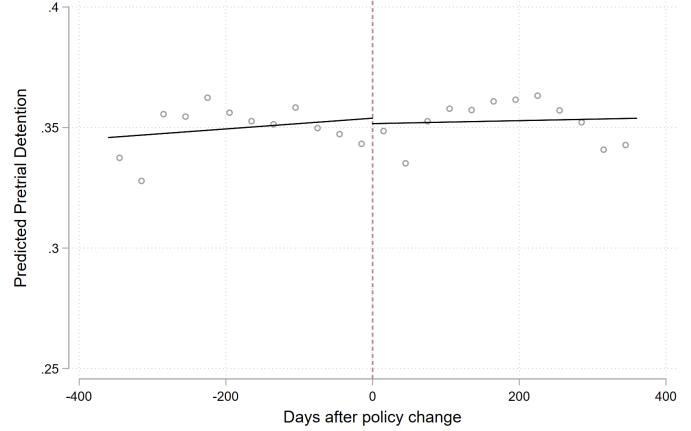


Notes: These figures plot tests of the regression discontinuity design. Each figure plots linear fits of the outcome listed and means of the outcome variable in 30 day bins.

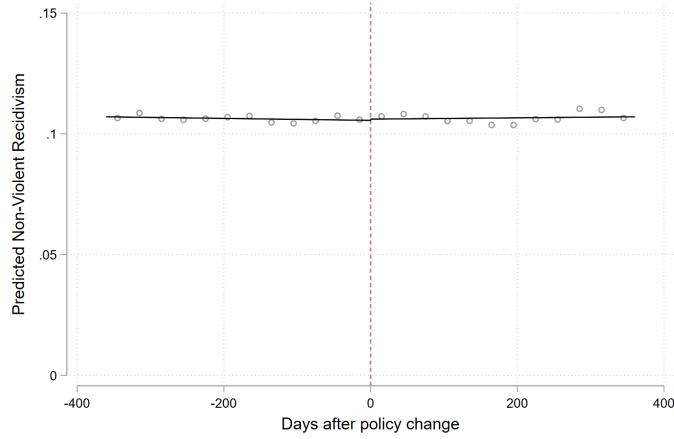
Figure 4: Regression Discontinuity Results for Predicted Values



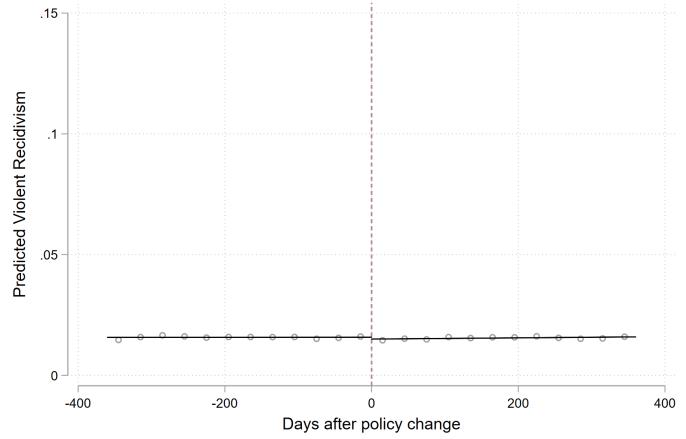
(a) Predicted Probability of Release on Non-financial Bond



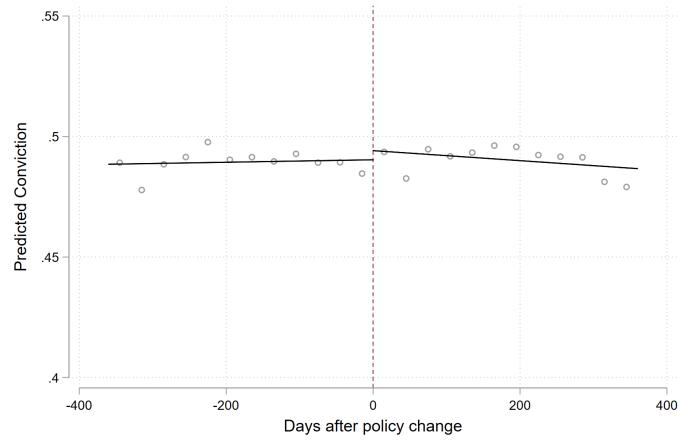
(b) Predicted Probability of Pretrial Detention



(c) Predicted Probability of Non-Violent Recidivism



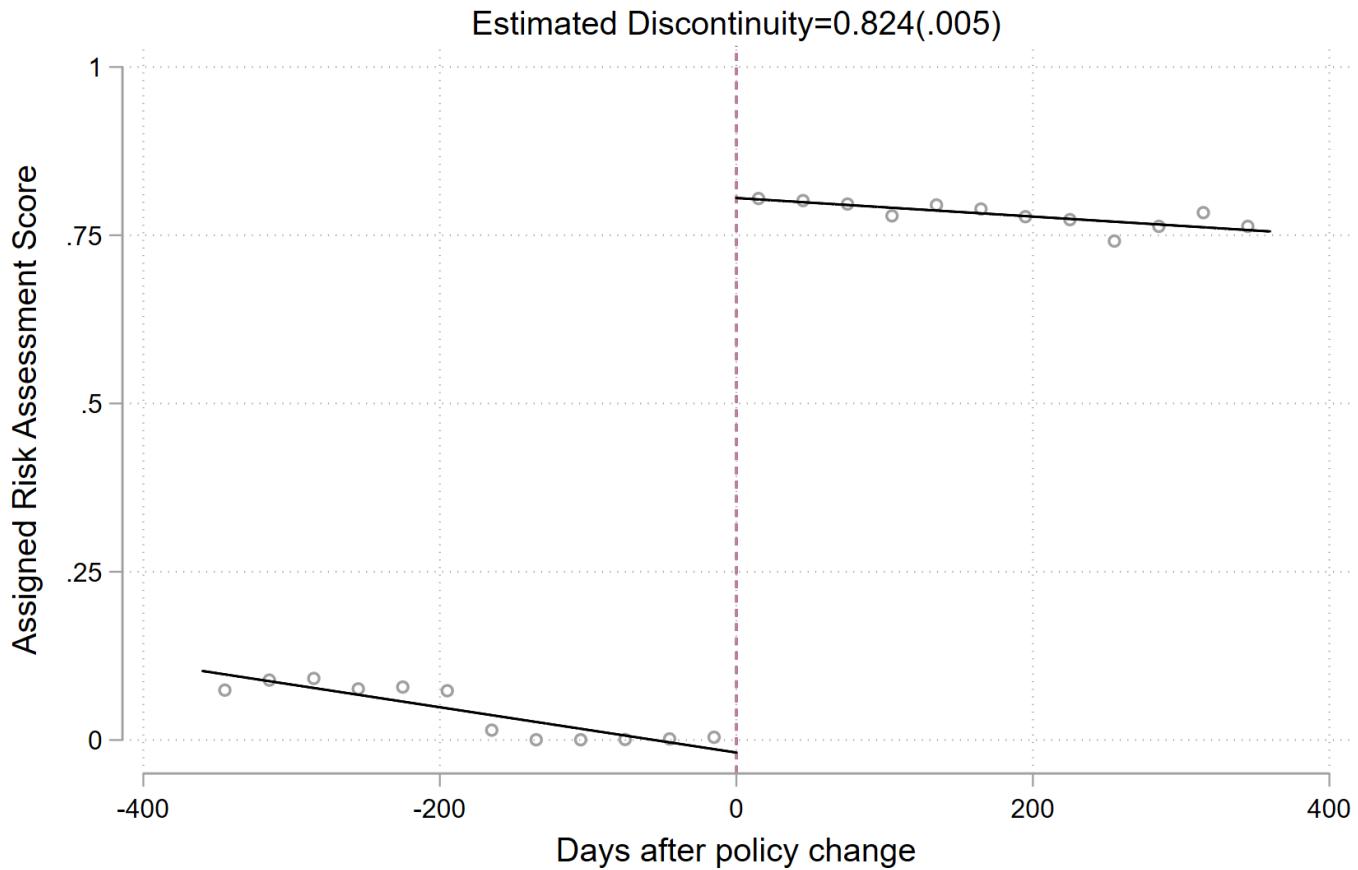
(d) Predicted Probability of Violent Recidivism



(e) Predicted Probability of Conviction

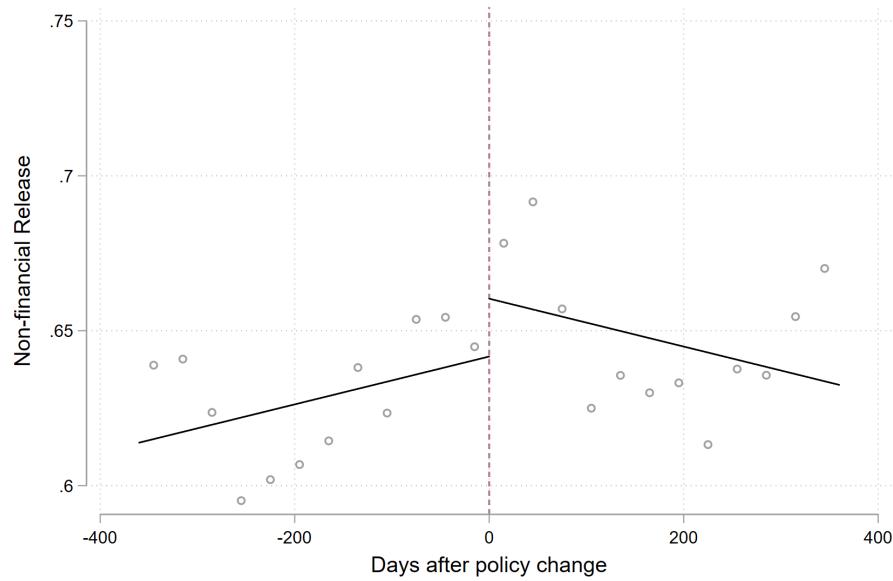
Notes: These figures plot tests of the regression discontinuity design. Each figure plots linear fits of the outcome listed and means of the outcome variable in 30 day bins. Outcome variables are predicted using observable case and defendant characteristics. Specifically, we use race, age, gender, criminal 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

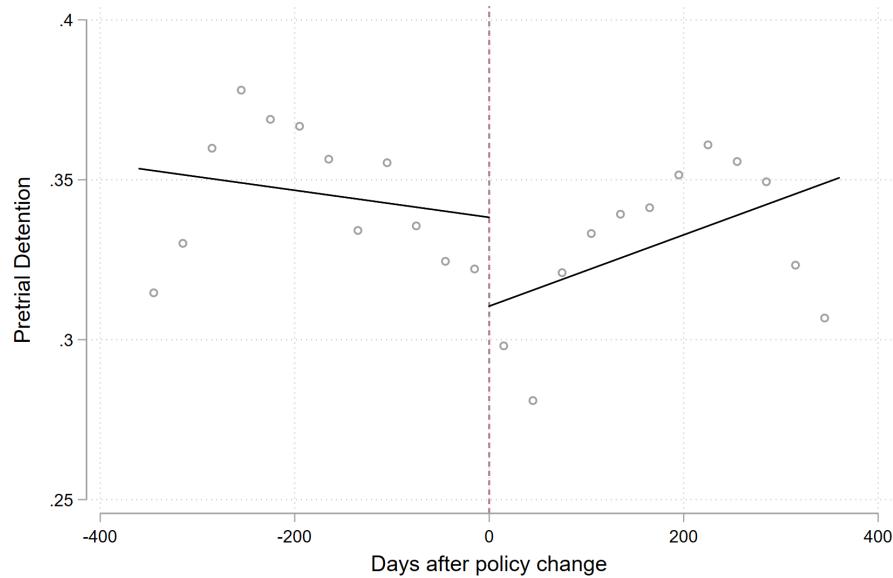


Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on score usage by plotting the mean of risk assessment score take-up in 30 day bins with linear fits. The outcome variable takes on a value of one if a defendant has a risk assessment 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.

Figure 6: Regression Discontinuity Results for Non-financial Bond and Pretrial Detention



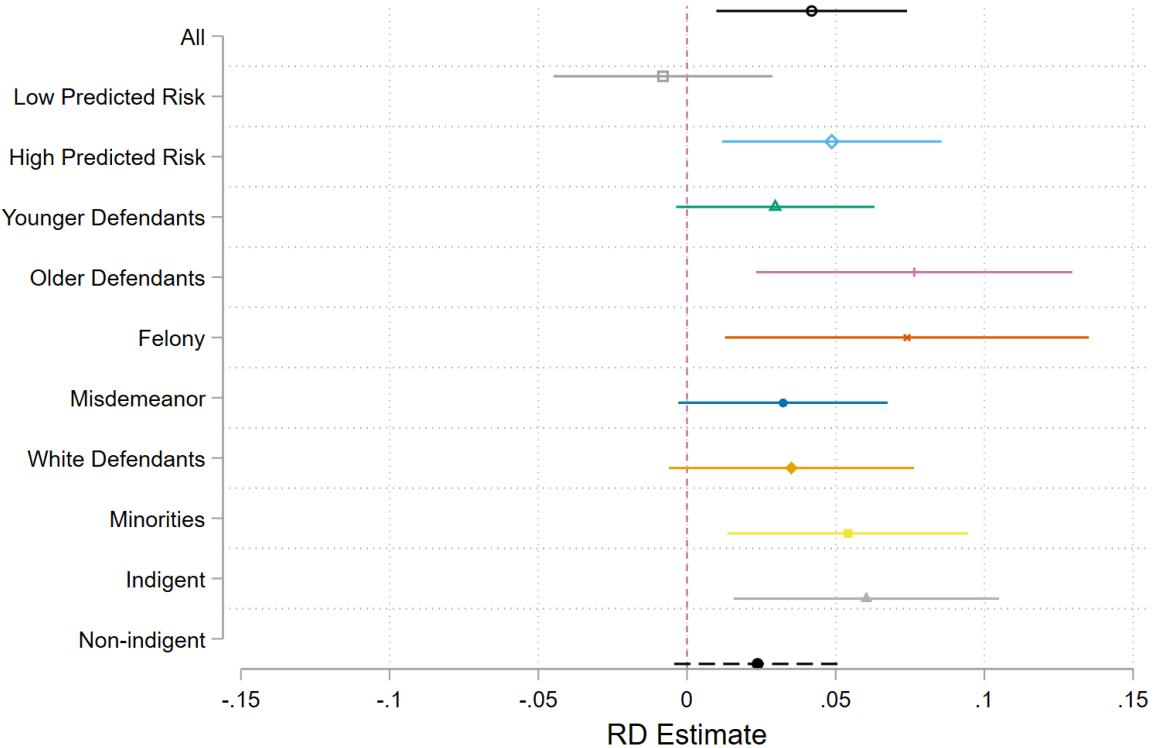
(a) Non-financial Bond



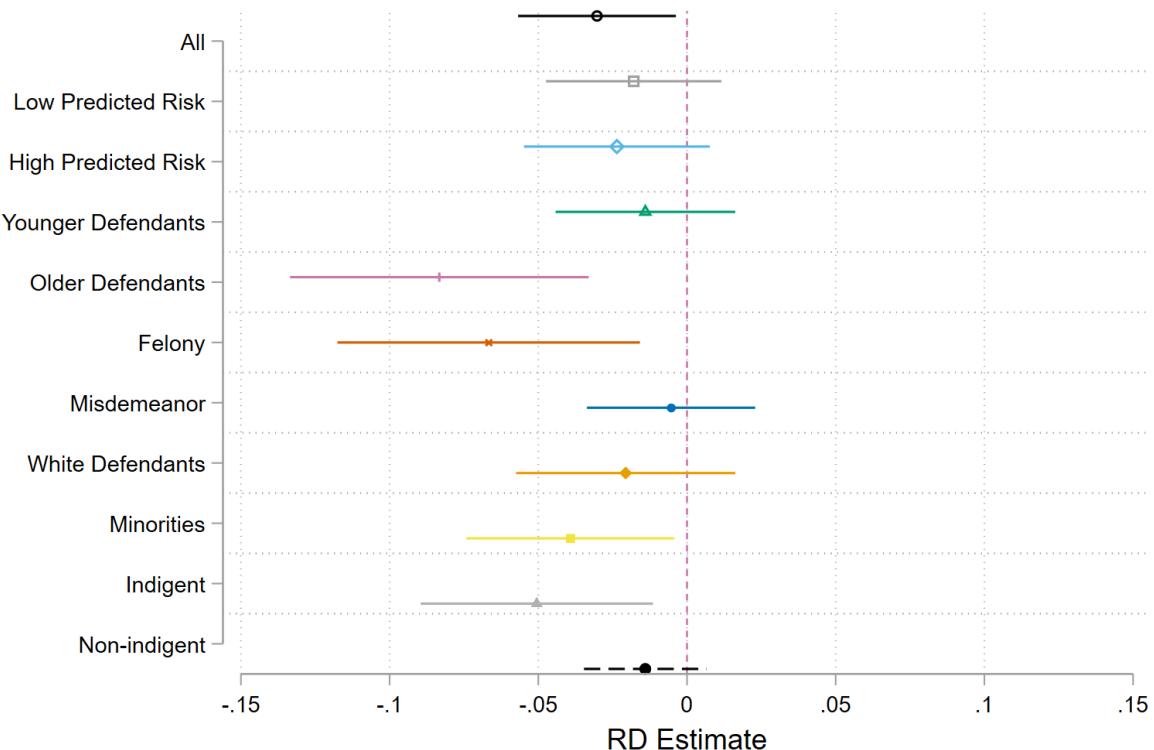
(b) Pretrial Detention

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 30 day bins with linear fits. A bandwidth of 360 days is shown.

Figure 7: The Effect of Risk Assessment Scores on Non-financial Release by Subgroup



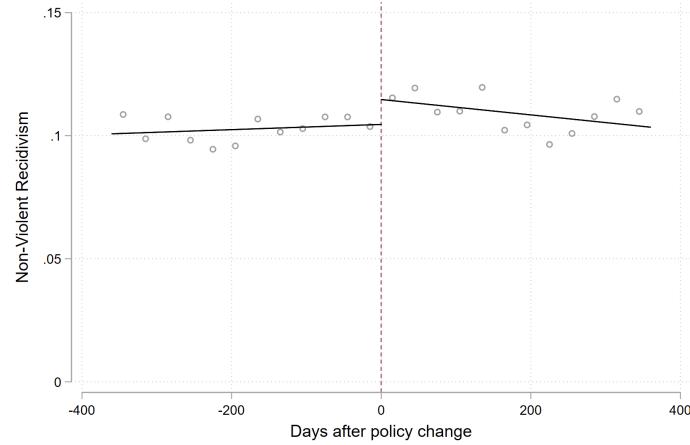
(a) Non-financial bond



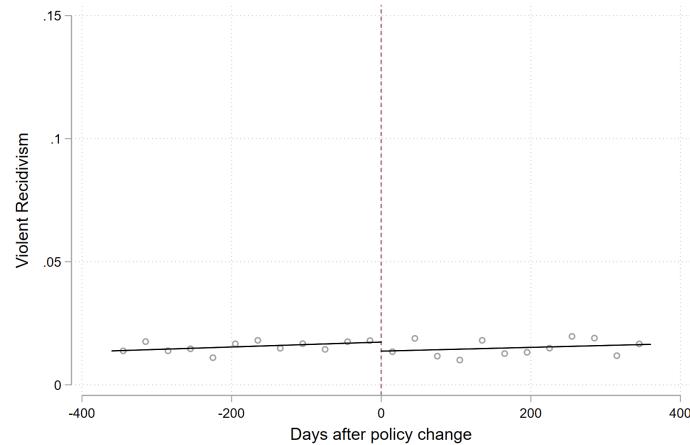
(b) Pretrial Detention

Each figure presents regression discontinuity results from eleven separate regressions. Each estimate is calculated using the MSE optimal bandwidth. Risk is predicted using observable case and defendant characteristics (race, indigent status, age, criminal history, gender, mental health, citizenship, misdemeanor, day of the week and court). Young defendants are younger than the average age (33 years). Older defendants are older than the average age. Non-minorities are only white defendants. Minorities are non-white or Hispanic.

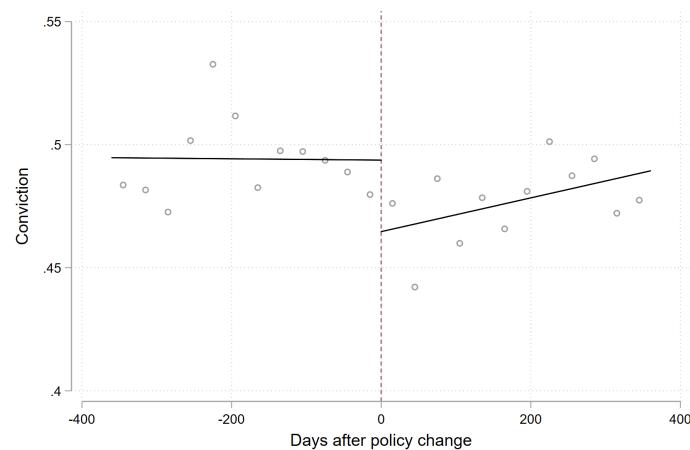
Figure 8: Regression Discontinuity Results for Conviction and Pretrial Crime



(a) Probability of Non-Violent Pretrial Crime



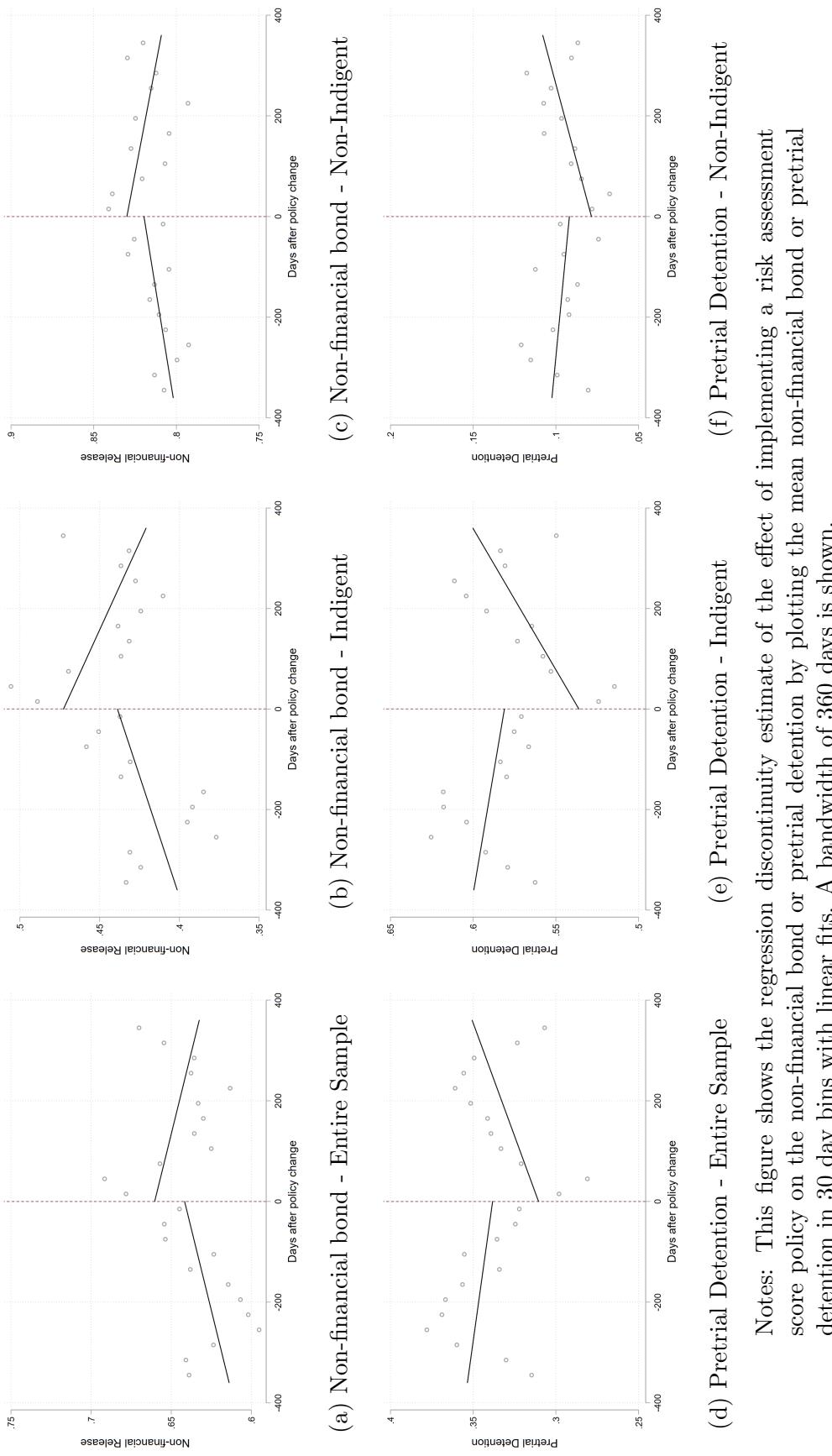
(b) Probability of Violent Pretrial Crime



(c) Probability of Conviction

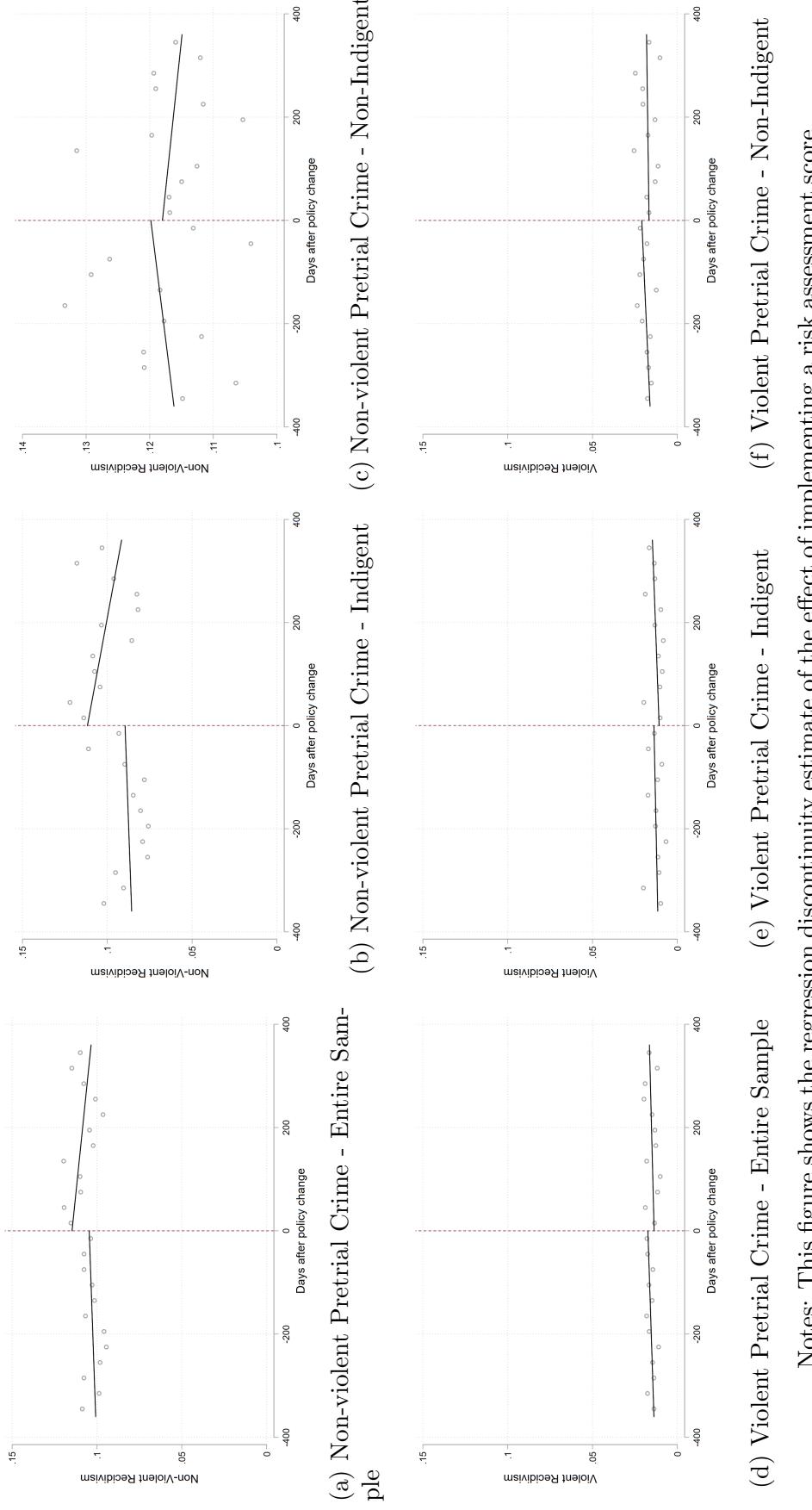
Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on conviction, non-violent and violent pretrial crime by plotting the mean of the outcome variable in 30 day bins with linear fits. A bandwidth of 360 days is shown.

Figure 9: Indigent Regression Discontinuity Results for Non-financial Bond and Pretrial Detention



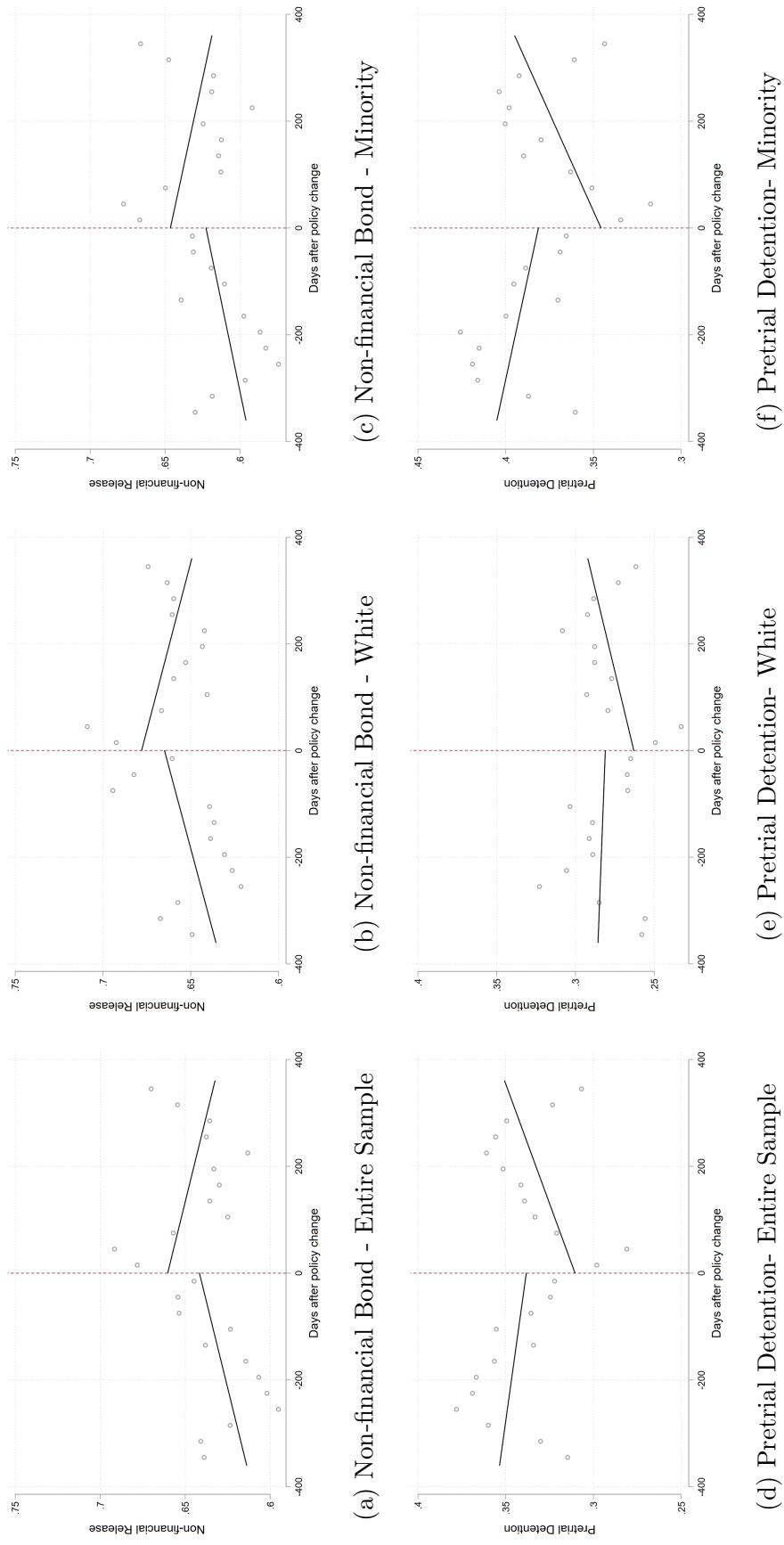
Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on the non-financial bond or pretrial detention by plotting the mean non-financial bond or pretrial detention in 30 day bins with linear fits. A bandwidth of 360 days is shown.

Figure 10: Indigent Regression Discontinuity Results for Pretrial Crime



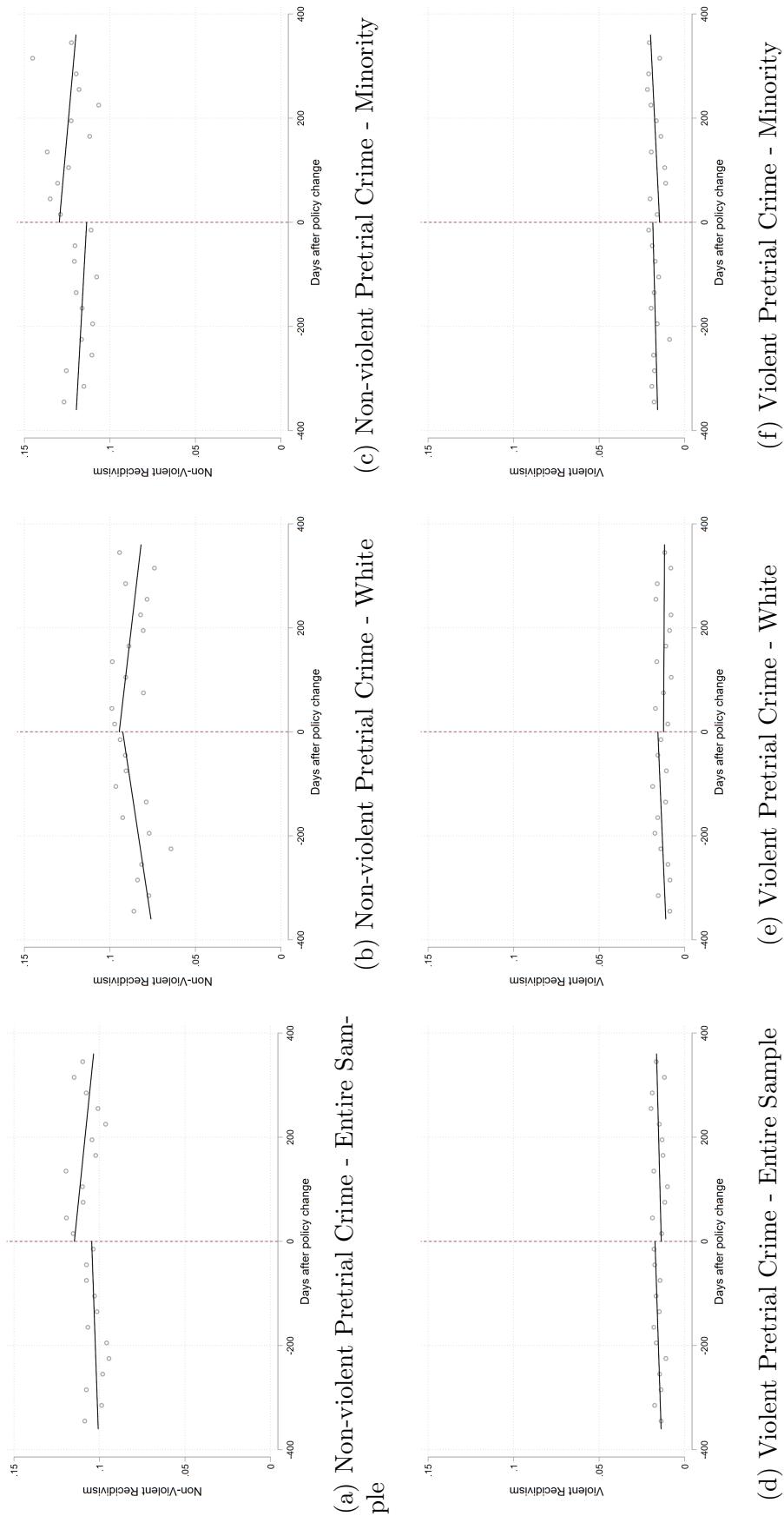
Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on pretrial crime by plotting the mean non-financial bond or pretrial detention in 30 day bins with linear fits. A bandwidth of 360 days is shown.

Figure 11: Regression Discontinuity Results for Non-financial Bond and Pretrial Detention by Race



Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on the non-financial bond or pretrial detention by plotting the mean non-financial bond or pretrial detention in 30 day bins with linear fits. A bandwidth of 360 days is shown. White defendants are only white. Minority defendants are Hispanic or non-white.

Figure 12: Regression Discontinuity Results for Pretrial Crime by Race



Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on pretrial crime by plotting the mean non-financial bond or pretrial detention in 30 day bins with linear fits. A bandwidth of 360 days is shown. White defendants are only white. Minority defendants are Hispanic or non-white.

Table 1: Summary Statistics

	Mean	Standard Deviation	Number of Observations
Case Characteristics			
White Defendant	0.42	0.49	143,092
Minority Defendant	0.58	0.49	143,092
Misdemeanor	0.6772	0.468	143,092
Defendant Age	32.5302	11.220	143,089
United States Citizen	0.8941	0.308	143,092
Male	0.7578	0.428	143,077
Indigent	0.5124	0.500	143,092
No Prior Offenses	0.7638	0.425	143,092
No Mental Health Flag	0.8762	0.329	143,092
Outcomes			
Non-financial Release	0.6253	0.484	127,904
Pretrial Detention	0.3532	0.478	143,092
Violent Pretrial Crime	0.0154	0.123	143,092
Non-Violent Pretrial Crime	0.1056	0.307	143,092
Conviction	0.4904	0.500	143,092

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. Defendants are white if they are white and not Hispanic. Minority defendants are either non-white or Hispanic

Table 2: Release Regression Discontinuity Results

	<i>2x Optimal Bandwidth</i>	<i>1.5x Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>			
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome: Non-financial Bail						
RD_Estimate	0.0285*** (0.0101)	0.0320*** (0.0113)	0.0406*** (0.0117)	0.0433*** (0.0130)	0.0411** (0.0160)	0.0423*** (0.0159)
Observations	25742	47484	19136	36052	12572	23440
FDR q-value	0.009	0.013	0.002	0.005	0.04	0.04
Bandwidth	185.0	337.6	138.8	253.2	92.51	168.8
Outcome: Pretrial Detention						
RD_Estimate	-0.0292*** (0.00731)	-0.0322*** (0.0102)	-0.0307*** (0.00844)	-0.0332*** (0.0117)	-0.0310** (0.0129)	-0.0255* (0.0144)
Observations	39354	45418	28734	34086	18862	22300
FDR q-value	0.001	0.01	0.001	0.013	0.04	0.19
Bandwidth	247.5	288.1	185.6	216.0	123.7	144.0
Controls	Y	Y	Y	Y	Y	Y
Quadratic	N	Y	N	Y	N	Y
Running Variable Control	Y	Y	Y	Y	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

	<i>2x Optimal Bandwidth</i>	<i>1.5x Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>			
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome: Non-Violent Pretrial Crime						
RD Estimate	0.00957*	0.0116*	0.00979*	0.0101	0.0108	0.00985
	(0.00507)	(0.00677)	(0.00586)	(0.00782)	(0.00727)	(0.00951)
Observations	58646	73300	44142	55922	28924	37706
FDR q-value	0.059	0.109	0.119	0.324	0.17	0.502
Bandwidth	372.1	474.6	279.1	356.0	186.0	237.3
Outcome: Violent Pretrial Crime						
RD Estimate	-0.00474**	-0.00399	-0.00239	-0.00327	-0.00319	-0.000657
	(0.00224)	(0.00273)	(0.00259)	(0.00316)	(0.00321)	(0.00386)
Observations	47280	71312	35572	54348	23258	36462
FDR q-value	0.043	0.143	0.355	0.367	0.322	0.865
Bandwidth	300.6	460.4	225.4	345.3	150.3	230.2
Outcome: Conviction						
RD Estimate	-0.0322***	-0.0232**	-0.0289***	-0.0112	-0.0202*	-0.0105
	(0.00761)	(0.0108)	(0.00879)	(0.0124)	(0.0113)	(0.0152)
Observations	59880	66434	44996	50492	29616	33590
FDR q-value	0.001	0.054	0.002	0.367	0.125	0.61
Bandwidth	381.0	427.8	285.8	320.9	190.5	213.9
Controls	Y	Y	Y	Y	Y	Y
Quadratic	N	Y	N	Y	N	Y
Running Variable Control	Y	Y	Y	Y	Y	Y

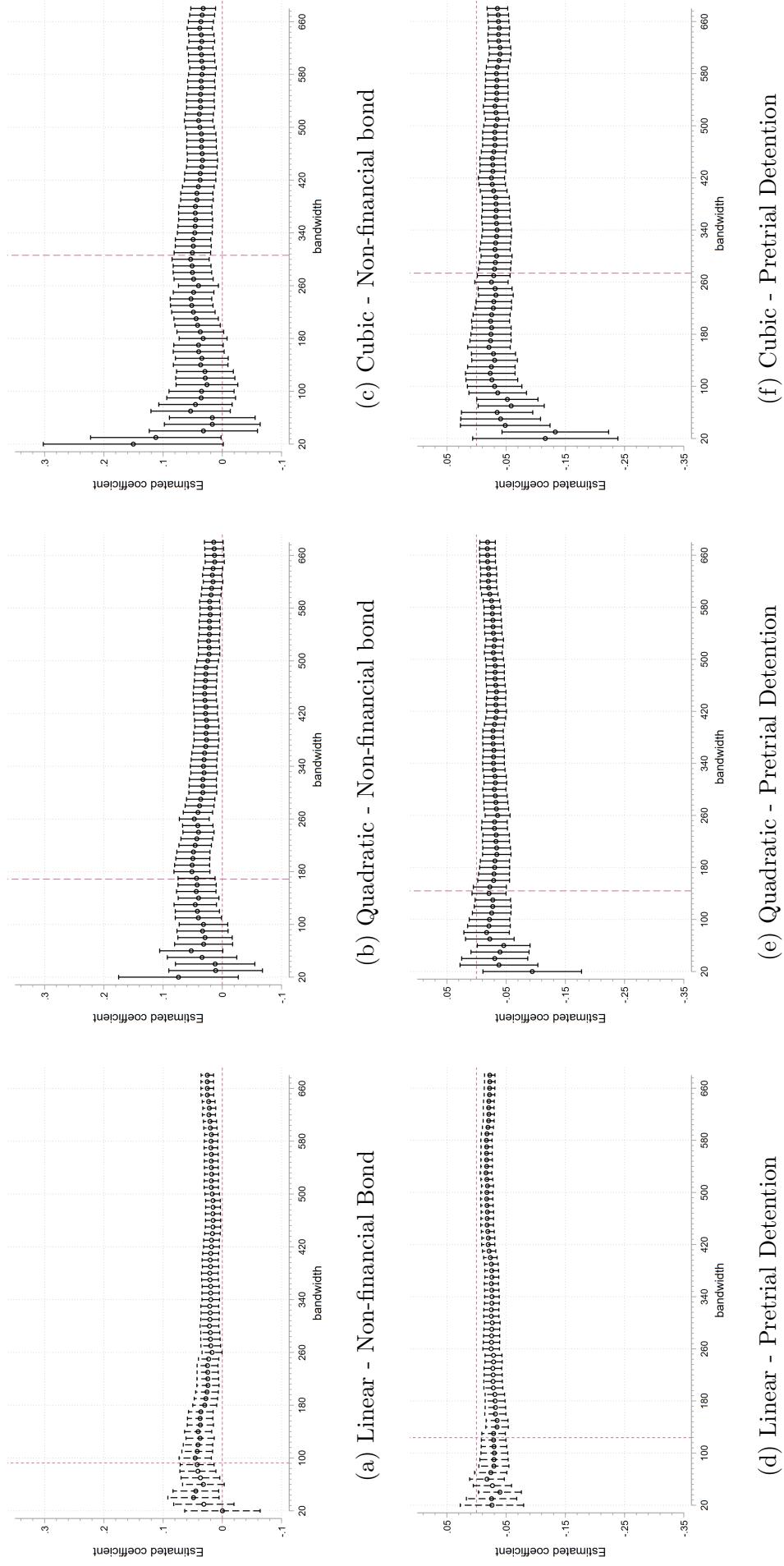
Standard errors in parentheses

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

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.

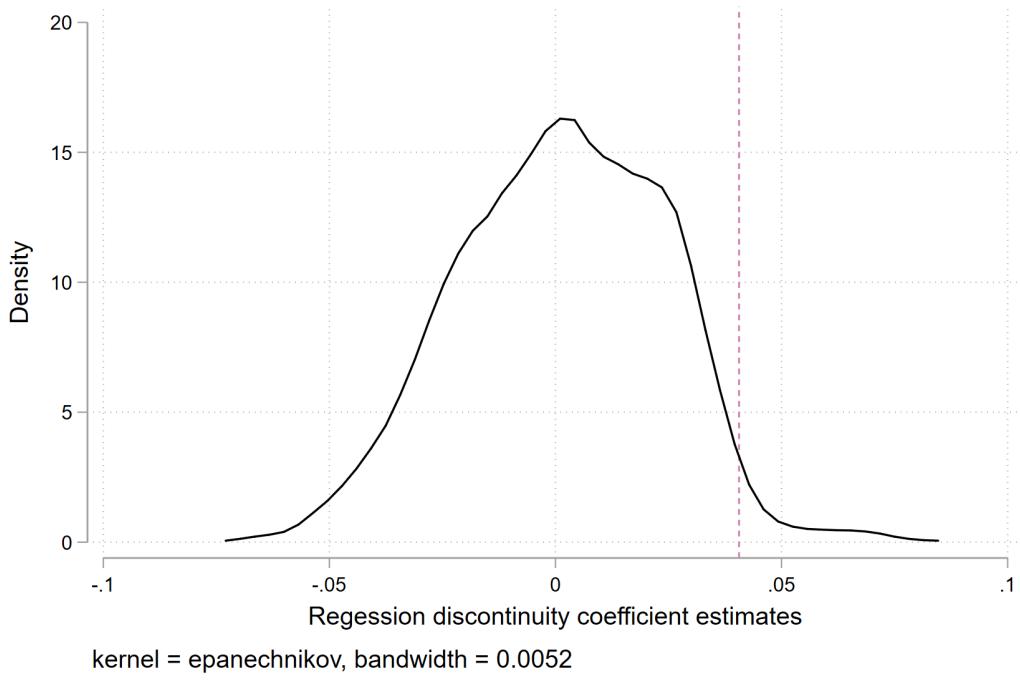
A Appendix

Figure A1: Non-financial and Pretrial detention Robustness

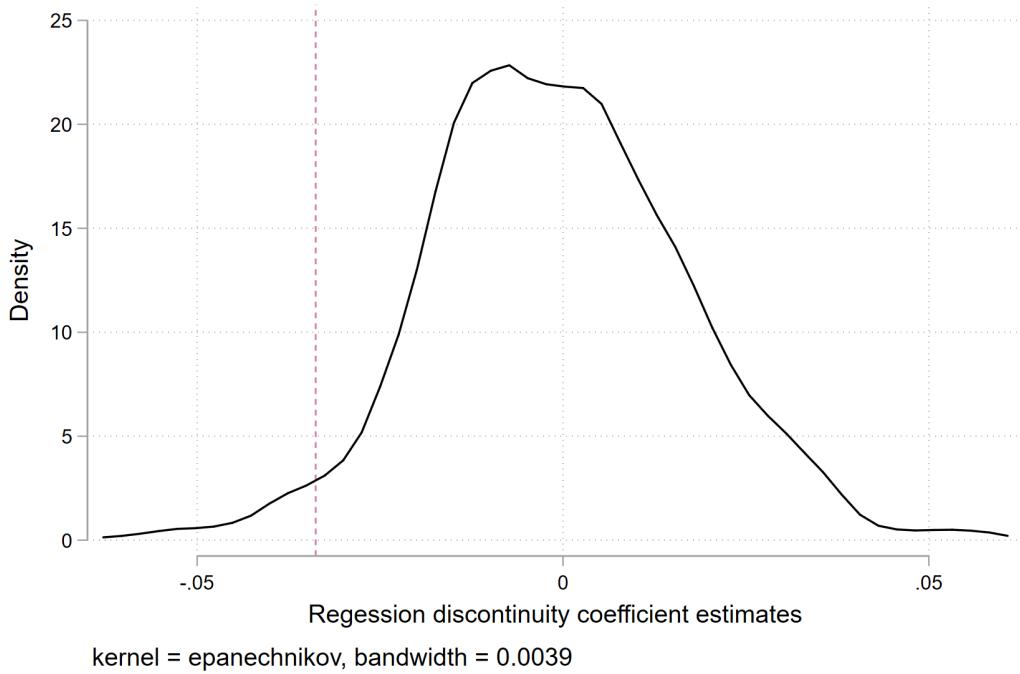


Notes: Each figure plots coefficients from 64 different regressions using different bandwidths. Ninety-five percent confidence intervals are also presented. The optimal MSE bandwidth for each specification is marked with the dashed line.

Figure A2: Reassigning Treatment Date



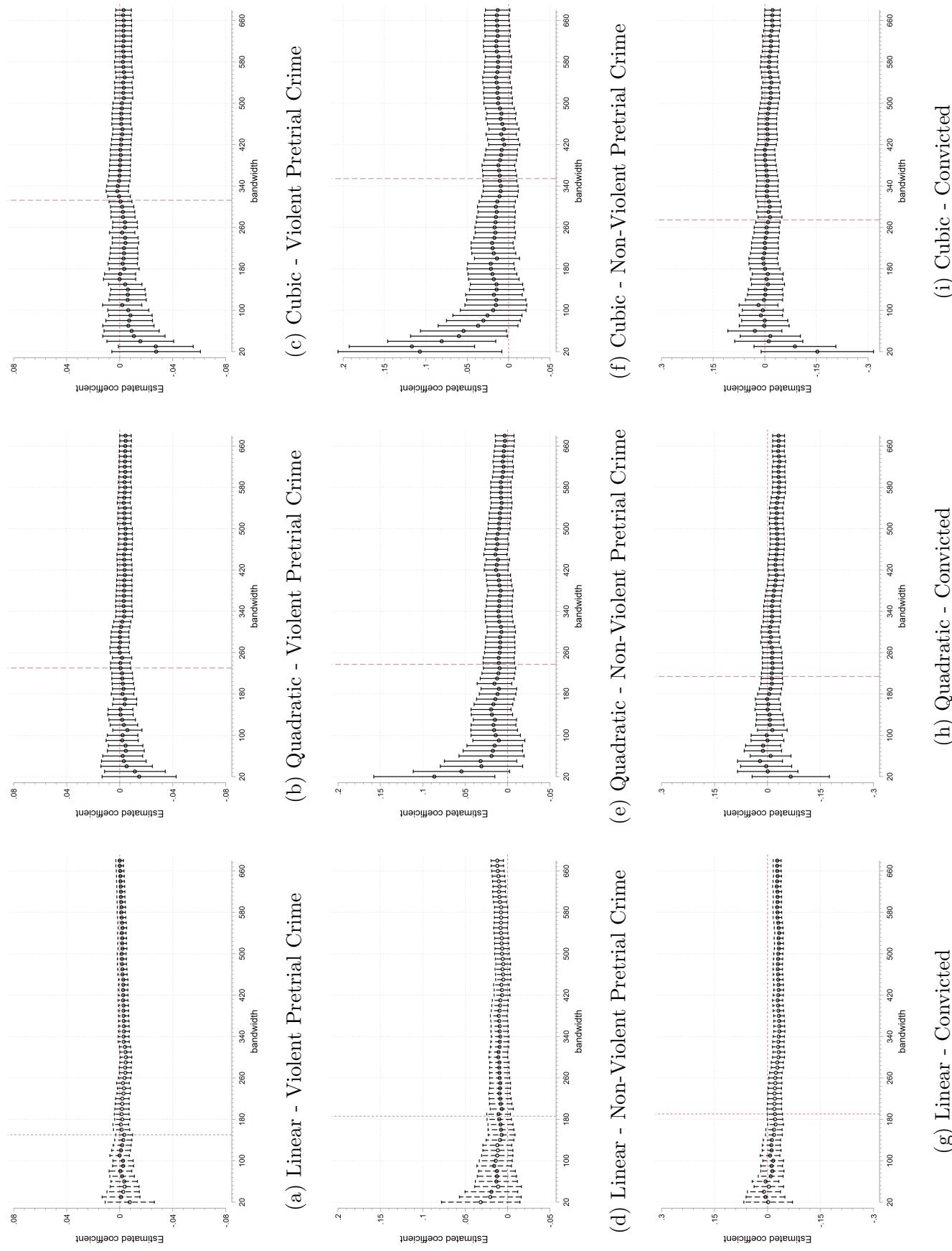
(a) Outcome: Release on Non-financial Bond



(b) Outcome: Pretrial Detention

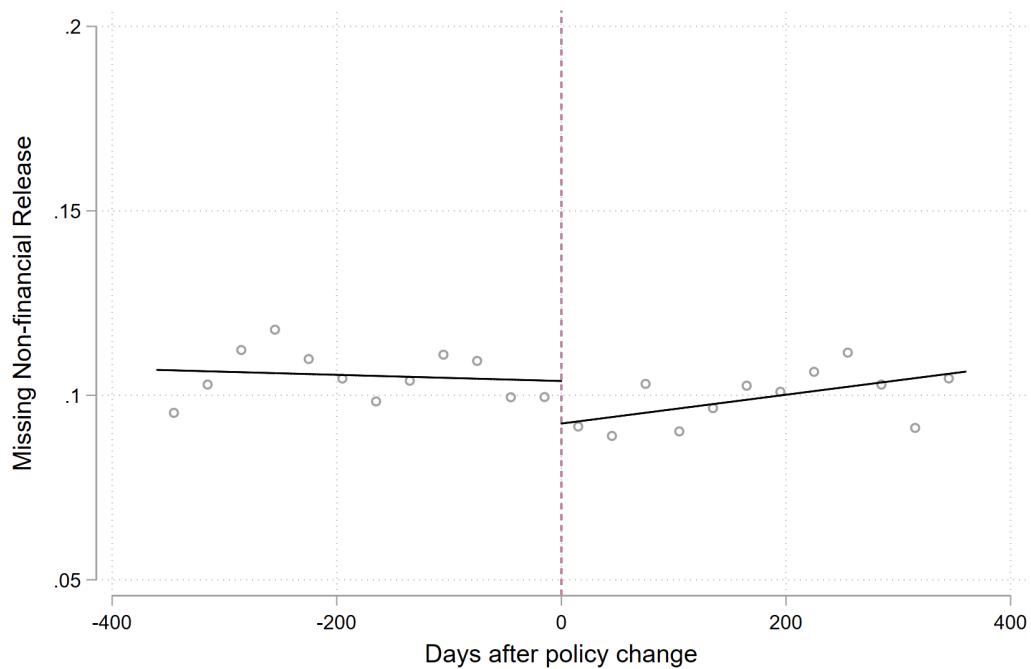
Notes: This figure plots the distribution of 910 regression discontinuity coefficients from equation (1) using pre-treatment data. Dashed lines are treatment effects from Table 2. For the probability of release on non-financial bond, our estimate reported in Table 2 is greater than 97.99 percent of all placebo estimates. For pretrial detention, our estimate reported in Table 2 is less than 95.88 percent of all placebo estimates

Figure A3: Pretrial Crime and Conviction Robustness



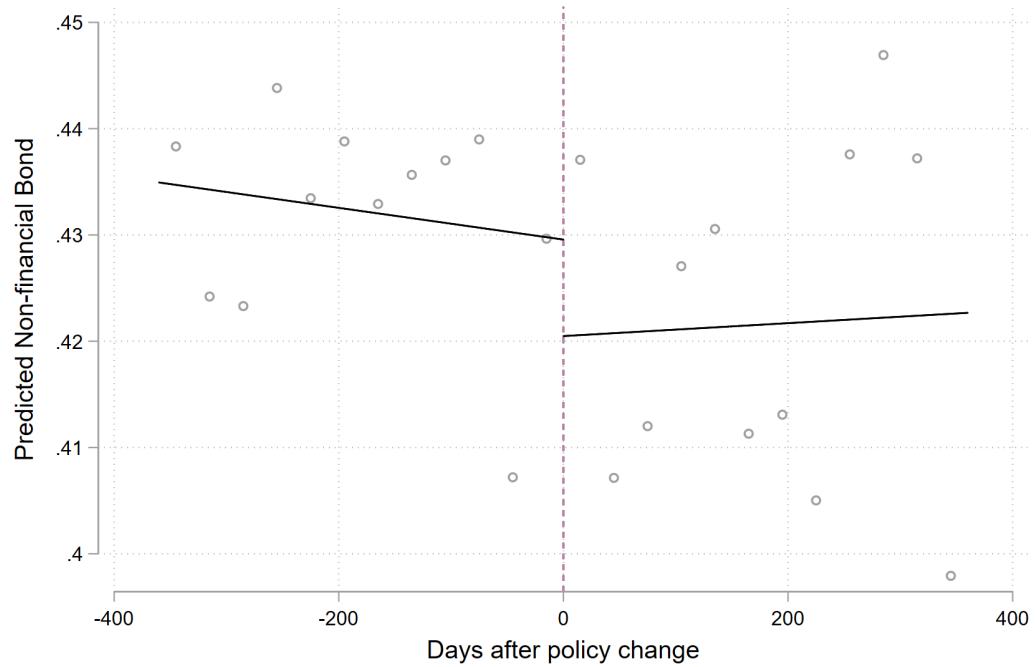
Notes: Each figure plots coefficients from 64 different regressions using different bandwidths. 95 percent confidence intervals are also presented. The optimal MSE bandwidth for each specification is marked with the dash line.

Figure A4: Regression Discontinuity Results for the Probability of Missing Outcome Data



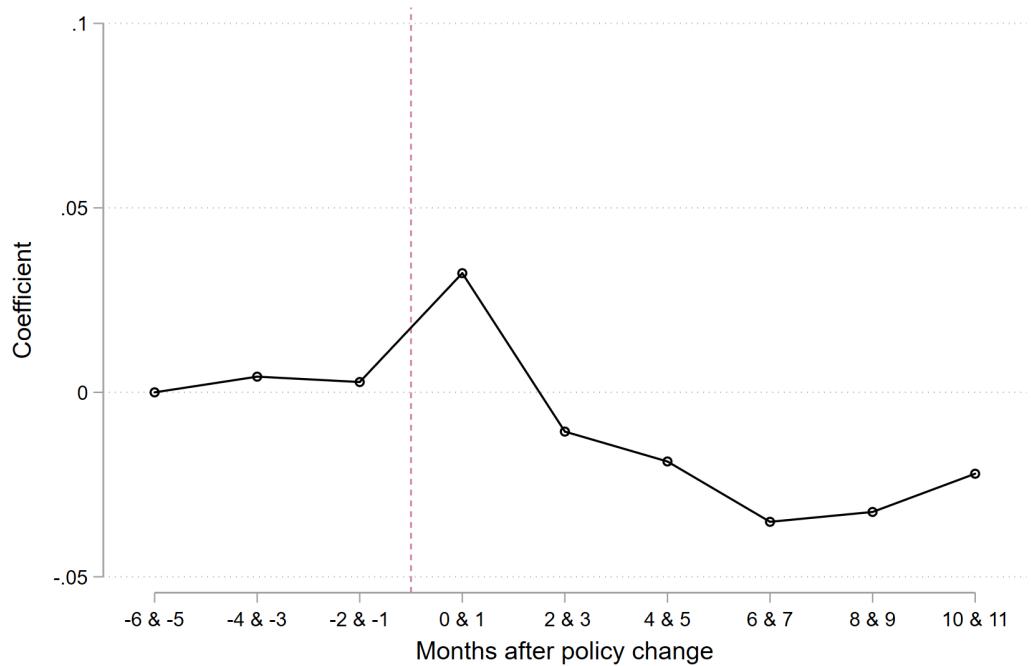
Notes: This figure shows the regression discontinuity estimate of the effect of implementing a risk assessment score policy on the likelihood of missing data by plotting the mean of the probability of missing in 30 day bins with linear fits. A bandwidth of 360 days is shown.

Figure A5: Regression Discontinuity Results for Predicted Probability of Release on non-financial bond for Defendants with Missing Outcome Data



Notes: This figure plots an additional test of the regression discontinuity design. This graph includes linear fits of the predicted probability of release on non-financial bond and means of the predicted probability of release on non-financial bond in 30 day bins. Outcome variables are predicted using observable case and defendant characteristics. A bandwidth of 360 days is shown. The RD is calculated only using observations from defendants who are missing data on non-financial bond.

Figure A6: Dynamic Effects of Risk Assessment Scores



Notes: This figure plots the coefficients from the regression of non-financial bond on indicators for months before or after risk assessment adoption. 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. A court-specific time trend is also included.

Table A1: Tests of the identifying assumption of the RD analysis

	Court 0	Court 3	Court 4	Court 5	Court 6	Court 7	Court 8	Court 9	Court 10	Court 11	Court 12
RD_Estimate	-0.00720 (0.00771)	0.00723 (0.00697)	0.000311 (0.00361)	-0.00361 (0.00442)	-0.000724 (0.00418)	0.00118 (0.00762)	0.000806 (0.00358)	0.000363 (0.00411)	0.00454 (0.00744)	0.00174 (0.00380)	-0.0137* (0.00823)
Observations	30638	37860	39698	23258	28568	31322	38320	28734	33238	35572	22300
Bandwidth	196.6	238.4	250.0	150.1	184.1	200.7	241.5	185.5	211.4	225.3	144.5
Run. Var. Control	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

	White Defendants	Minority Defendant	Misdemeanor	Def. Age	US Citizen	Male	Indigent	No Priors	Mental Health
RD_Estimate	-0.0114 (0.0123)	0.0114 (0.0123)	0.0155 (0.0116)	0.704** (0.336)	0.00114 (0.00776)	-0.00909 (0.00990)	0.00138 (0.0141)	-0.00992 (0.0109)	0.00425 (0.00698)
Observations	25162	25162	24712	16818	25162	29286	19190	25708	34362
Bandwidth	162.4	162.4	159.5	110.8	162.4	188.6	125.0	165.2	218.5
Run. Var. Control	Y	Y	Y	Y	Y	Y	Y	Y	Y

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

Notes: Each cell represents results for separate regressions. Robust standard errors are in parentheses. All specifications control for a linear function of distance from policy enactment in which the slope is allowed to vary on either side of the cutoff. The optimal (MSE) bandwidth is used to determine the sample for each separate regression. Some courts (1,2,13) do not occur enough times in our sample to estimate effects.

Table A2: Regression Discontinuity Results for Predicted Outcomes

	<i>Optimal Bandwidth</i>	
	(1)	(2)
Outcome: Predicted Non-financial Bail		
RD_Estimate	-0.00153 (0.00673)	0.00357 (0.00753)
Observations	19604	36288
Bandwidth	128.3	229.3
Outcome: Predicted Pretrial Detention		
RD_Estimate	-0.000287 (0.00839)	-0.00345 (0.00928)
Observations	19604	37202
Bandwidth	128.2	234.9
Outcome: Predicted Non-Violent Pretrial Crime		
RD_Estimate	0.00113 (0.00152)	0.00130 (0.00171)
Observations	20440	37556
Bandwidth	133.7	236.5
Outcome: Predicted Violent Pretrial Crime		
RD_Estimate	-0.00101 (0.000630)	-0.00117 (0.000722)
Observations	23070	41080
Bandwidth	149.7	258.1
Outcome: Predicted Conviction		
RD_Estimate	0.00376 (0.00487)	0.00627 (0.00578)
Observations	20266	33238
Bandwidth	132.9	211.3
Controls	N	N
Quadratic	N	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. 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 effect. Column (1) presents a linear functional form and column (2) is quadratic.

Table A3: Release Regression Discontinuity Results for Indigent and Non-Indigent Defendants

	<i>2x Optimal Bandwidth</i>	<i>1.5x Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>		
	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Indigent</i>					
Outcome: Non-financial Bail					
RD_Estimate	0.0313*	0.0400*	0.0538***	0.0544**	0.0604**
	(0.0172)	(0.0206)	(0.0199)	(0.0237)	(0.0245)
Observations	12794	20578	9438	15534	6120
Bandwidth	195.1	309.7	146.3	232.3	97.54
Outcome: Pretrial Detention					
RD_Estimate	-0.0428***	-0.0418**	-0.0383**	-0.0480**	-0.0466**
	(0.0136)	(0.0191)	(0.0158)	(0.0220)	(0.0194)
Observations	20802	23886	15106	17812	9844
Bandwidth	257.2	298.2	192.9	223.6	128.6
<i>Panel B: Non-Indigent</i>					
Outcome: Non-financial Bail					
RD_Estimate	0.0153	0.0110	0.0202*	0.0198	0.0262*
	(0.0104)	(0.0113)	(0.0118)	(0.0130)	(0.0144)
Observations	20254	37668	15112	28618	9940
Bandwidth	270.4	511.2	202.8	383.4	135.2
Outcome: Pretrial Detention					
RD_Estimate	-0.0145*	-0.0206**	-0.0257***	-0.0223**	-0.0145
	(0.00755)	(0.00896)	(0.00863)	(0.0103)	(0.0105)
Observations	20392	31924	15118	24010	9896
Bandwidth	261.2	412.6	195.9	309.5	130.6
Controls	Y	Y	Y	Y	Y
Quadratic	N	Y	N	Y	N
Running Variable Control	Y	Y	Y	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 indigent and non-indigent subgroups respectively.

Table A4: Pretrial Crime Regression Discontinuity Results for Indigent and Non-Indigent Defendants

	<i>2x Optimal Bandwidth</i>	<i>1.5x Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>			
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Indigent Defendants						
Outcome: Non-Violent Pretrial Crime						
RD_Estimate	0.0202** (0.00827)	0.0217** (0.0100)	0.0142 (0.00962)	0.0144 (0.0116)	0.0152 (0.0119)	0.0182 (0.0141)
Observations	22040	33760	16220	25796	10550	17114
Bandwidth	273.4	431.4	205.1	323.5	136.7	215.7
Outcome: Violent Pretrial Crime						
RD_Estimate	-0.00216 (0.00249)	-0.00298 (0.00297)	-0.00341 (0.00293)	-0.00217 (0.00349)	0.000587 (0.00360)	-0.00151 (0.00431)
Observations	31534	52318	24042	38294	15784	26038
Bandwidth	400.3	653.9	300.2	490.4	200.2	326.9
Panel B: Non-Indigent						
Outcome: Non-Violent Pretrial Crime						
RD_Estimate	-0.00154 (0.00767)	-0.00166 (0.00900)	0.000579 (0.00883)	0.00599 (0.0104)	0.00379 (0.0107)	0.00561 (0.0127)
Observations	27180	43276	20392	32626	13534	21884
Bandwidth	348.3	564.4	261.2	423.3	174.1	282.2
Outcome: Violent Pretrial Crime						
RD_Estimate	-0.00632* (0.00326)	-0.00571 (0.00411)	-0.00428 (0.00376)	-0.00501 (0.00474)	-0.00475 (0.00466)	-0.00386 (0.00584)
Observations	24540	34648	18580	26308	12300	17586
Bandwidth	316.1	450.9	237.1	338.2	158.0	225.5
Controls	Y	Y	Y	Y	Y	Y
Quadratic	N	Y	N	Y	N	Y
Running Variable Control	Y	Y	Y	Y	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 indigent and non-indigent subgroups respectively.

Table A5: Release Regression Discontinuity Results for White and Minority Defendants

	<i>2x Optimal Bandwidth</i>	<i>1.5x Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>		
	(1)	(2)	(3)	(4)	(5)
Panel A: White Defendants					
Outcome: Non-financial Bail					
RD_Estimate	0.0181 (0.0126)	0.0254 (0.0202)	0.0145 (0.0146)	0.0191 (0.0164)	0.0286 (0.0178)
Observations	15946	13968	11694	20898	7818
Bandwidth	250.5	221.0	187.8	331.5	125.2
Outcome: Pretrial Detention					
RD_Estimate	-0.0202** (0.00981)	-0.0242* (0.0123)	-0.0234** (0.0112)	-0.0252* (0.0143)	-0.0259* (0.0138)
Observations	18118	25624	13384	19128	8878
Bandwidth	266.5	378.4	199.8	283.8	133.2
Panel B: Minority Defendants					
Outcome: Non-financial Bail					
RD_Estimate	0.0316** (0.0128)	0.0409*** (0.0154)	0.0540*** (0.0147)	0.0609*** (0.0178)	0.0526*** (0.0181)
Observations	17240	26618	12610	20178	8240
Bandwidth	221.3	342.6	165.9	257.0	110.6
Outcome: Pretrial Detention					
RD_Estimate	-0.0336*** (0.0102)	-0.0362** (0.0142)	-0.0354*** (0.0118)	-0.0389** (0.0163)	-0.0306** (0.0144)
Observations	22662	26158	16476	19470	10746
Bandwidth	249.2	290.7	186.9	218.0	124.6
Controls	Y	Y	Y	Y	Y
Quadratic	N	Y	N	Y	N
Running Variable Control	Y	Y	Y	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 A6: Pretrial Crime Regression Discontinuity Results for White and Minority Defendants

	<i>2x Optimal Bandwidth</i>	<i>1.5x Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>			
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: White Defendants						
Outcome: Non-Violent Pretrial Crime						
RD_Estimate	0.00315 (0.00597)	-0.00201 (0.00864)	0.000673 (0.00694)	0.000465 (0.0100)	0.00195 (0.00856)	-0.00196 (0.0124)
Observations	35164	38926	26790	29422	18052	19820
Bandwidth	530.6	587.5	398.0	440.6	265.3	293.7
Outcome: Violent Pretrial Crime						
RD_Estimate	-0.00362 (0.00307)	-0.00231 (0.00367)	0.000665 (0.00354)	0.0000628 (0.00419)	-0.00116 (0.00431)	0.000121 (0.00513)
Observations	20734	31998	15640	24604	10390	16460
Bandwidth	306.7	483.0	230.0	362.2	153.4	241.5
Panel B: Minority Defendants						
Outcome: Non-Violent Pretrial Crime						
RD_Estimate	0.0135* (0.00737)	0.0209** (0.00965)	0.0134 (0.00845)	0.0196* (0.0111)	0.0160 (0.0103)	0.0163 (0.0135)
Observations	30596	39450	23212	30148	15026	20108
Bandwidth	340.2	447.4	255.1	335.6	170.1	223.7
Outcome: Violent Pretrial Crime						
RD_Estimate	-0.00527 (0.00323)	-0.00537 (0.00344)	-0.00441 (0.00372)	-0.00511 (0.00402)	-0.00556 (0.00466)	-0.00266 (0.00496)
Observations	25950	50834	19398	37754	12526	25630
Bandwidth	288.5	568.3	216.4	426.2	144.3	284.2
Controls	Y	Y	Y	Y	Y	Y
Quadratic	N	Y	N	Y	N	Y
Running Variable Control	Y	Y	Y	Y	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.

Table A7: Regression Discontinuity Results for the Probability of Missing Data

	$\mathcal{D}x$	<i>Optimal Bandwidth</i>	$1.5x$	<i>Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>	<i>Optimal Bandwidth</i>	<i>Optimal Bandwidth for Pr(Missing)</i>
Outcome: Missing Data							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
RD_Estimate	-0.00533 (0.00527)	0.000386 (0.00839)	-0.00693 (0.00614)	-0.00718 (0.00681)	-0.000794 (0.00703)	-0.00467 (0.00785)	-0.00557 (0.00955)
Observations	40600	34528	28734	53140	21372	40298	14016
Bandwidth	255.1	219.5	185.0	337.6	138.8	253.2	92.51
Controls	Y	Y	Y	Y	Y	Y	Y
Quadratic	N	Y	N	Y	N	Y	Y
Run. Var. Control	Y	Y	Y	Y	Y	Y	Y

Standard errors in parentheses

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

Notes: Each cell represents results for separate regression. Each column presents results for the probability of missing data for the outcome variable and the key independent variable is an indicator for policy enactment. Robust standard errors are in parentheses. All specifications control for a linear function of distance from policy enactment in which the slope is allowed to vary on either side of the cutoff. The optimal (MSE) bandwidth is used to determine the sample for each separate regression in the first three columns. Columns (1)-(6) use the optimal bandwidth determined in Table 2. Columns (7)-(8) use the optimal bandwidth for the Probability of Missing Data.