Do Mild Sentences Deter Crime?
Evidence using a Regression-Discontinuity Design

Björn Tyrefors Hinnerich,[♣] Per Pettersson-Lidbom,[♥] Mikael Priks[♠]

This version

September 9, 2016

Abstract

We study if harsher sentences deter or increase reoffending by exploiting two discontinuities in the Swedish legal system against driving under the influence. Above certain thresholds, individuals tend to be sentenced to prison, which in practice means one month in minimum-security institutions or electronic monitoring, as opposed to probation. The results show that individuals just above the thresholds commit fewer crimes (e.g., drug-related crimes and assaults) upon release than those receiving probation.

[♣] Department of Economics, Stockholm University, E-mail: bjorn.hinnerich@ne.su.se

[♥] Department of Economics, Stockholm University, E-mail: pp@ne.su.se

[♠] Department of Economics, Stockholm University, E-mail: mikael.priks@ne.su.se

1. Introduction

Imprisonment is highly debated both in the U.S. and in Europe due to the high monetary costs as well as ethical concerns. A potential benefit is that harsher penalties may reduce reoffending if the punishment in itself is considered sufficiently distasteful or if the treatment during incarceration is successful. Alternatively, harsher punishments may increase offending through peer effects or stigmatization. However, it has proven to be challenging to evaluate these conflicting hypotheses since (i) individuals sentenced to different punishments are not ex-ante comparable and (ii) sentencing takes on multiple "treatments" such as incarceration, probation and fines (e.g., Chalfin and McCrary, 2014 and Mueller-Smith, 2015). A recent literature attempts to address some of these problems, but the evidence is still limited when it comes to how mild and short detentions, such as minimum-security institutions or electronic monitoring, affect crime as compared to probation.

In this paper, we are able to shed light on this issue by exploiting two discontinuities in the Swedish punishments for driving under the influence. Specifically, above certain blood alcohol concentrations (BAC), individuals are often sentenced to imprisonment while those just below are sentenced to probation. In practice, imprisonment here implies either one month in a minimum-security institution, or one month of electronic monitoring. These two types of punishments restrict the personal freedom in similar ways and are mild in comparison to being sentenced to a prison with high security standard. Consequently, we can use a regression-discontinuity (RD) design to compare the reoffending behavior of ex-ante similar individuals that are sentenced to either probation or imprisonment. Our RD strategy is also *a priori* attractive since neither drivers nor the police can manipulate the BAC due the strict automated administrative procedures in Sweden. In other words, there is no reason to expect any sorting of observations around the BAC thresholds.[1]

The results indicate that the individuals just above the thresholds commit significantly fewer crimes (the reduced form effect is approximately 30 percent) after serving their time, compared to those sentenced to probation. The result seems to be due to deterrence as opposed to incapacitation since the effect increases over time. Another explanation would be that individuals receive rehabilitation while in prison and therefore commit fewer crimes in the future. Such possibilities are, however, highly limited for imprisonment with short duration.

Our results speaks to the literature on the optimal level of punishment. Previous empirical studies that exploit plausible exogenous sources of variation in punishments have so far showed

[1] We cannot reject the null hypothesis of no sorting by a formal statistical test (McCrary, 2008).

that milder forms of imprisonment tend to be better than harsher with respect to reoffending. For example, Chen and Shapiro (2007) show that inmates assigned to minimum-security institutions do not reoffend less, but possibly more, than those above a cut-off level for institutions with higher security. Mastrobuoni and Terlizzese (2015) provide evidence in favor of that one more year in a rehabilitating prison reduces reoffending compared to a year in a normal prison while Drago et al. (2011) suggest that harsh prison conditions increase post-release criminal activity.[2] Di Tella and Schargrodsky (2013) show that electronic monitoring reduces reoffending compared to imprisonment. These results call for the question if the punishment should be even lower than mild prison sentences and electronic monitoring. Our results suggest that, at least for drunk driving, mild imprisonment sentences including electronic monitoring are in fact superior to non-custodial sanctions. Thus, taken together with the results from the earlier literature, this indicates that mild prison sentences may therefore be optimal.

Our study is most closely related to Hansen (2013) who also uses a RD approach based on BAC thresholds. In contrast to our RD design, there are a much large number of discontinuities in various punishments at the thresholds, which makes it hard to estimate the effect of imprisonment on criminal behavior as noted by Hansen. He therefore draws the conclusion that harsher punishment, due to either imprisonment, probation or fees, or a combination of the different punishments, reduces crime.

Our paper is also related to the literature on specific deterrence, i.e., a convicted person's inclination to commit new crime. This literature uses other types of exogenous sources of variations in punishments, such as random allocations to judges or lawyers (e.g., Berube and Green, 2007, Green and Winik 2010, Abrams 2011, Aizer and Doyle, 2013 and Mueller-Smith 2015), randomized amnesties (e.g., van der Werff, 1979, Drago et al., 2009, Maurin and Ouss, 2009, Buonanno and Raphael, 2013, and Barbarino and Mastrobuani, 2014) and non-linearities in the judicial system (e.g., Kuziemko 2012 and Hjalmarsson 2009a). The overall result from this literature is inclusive.

Another related strand of literature focuses on general deterrence, i.e., that imprisonment may also deter the population at large. This literature includes the study of three strikes and you're out reform in California (Helland and Tabarrok, 2007), discontinuities at birthdays (Lee and McCrary, 2009, Hjalmarsson, 2009b and Hinnerich et al. 2016) and various law changes

[2] Katz el al. (2003), in contrast, show that prison death rates, which is used as a proxy for prison conditions, is negatively correlated with subsequent crime rates. Here, however, there is no exogenous variation in prison conditions.

(Kessler and Levitt, 1999, Raphael and Ludwig, 2003 and Owens, 2009). This literature also reveals mixed results.

The rest of paper is structured as follows. We discuss the background and the data in Section 2. Section 3 discusses the empirical design. The results are presented in Section 4 while Section 5 concludes.

2. Background and data

In this section, we describe the background and data. Starting with the Swedish laws against driving under the influence, there are essentially three different BAC thresholds: 0.2, 1.0 and 1.5, with significantly different levels of punishments. Individuals with a BAC level of above 0.2 only get a fine. In addition, they normally lose their driving license for one year.[3] In contrast, individuals with a BAC level above 1.0 may sometimes be imprisoned. They also lose their driving license (for two years) and are therefore required to redo the whole procedure of retaking the driving license. The harshest punishment is for individuals that have a BAC level above 1.5 since they are often sentenced to prison.[4] The length of imprisonment is normally one month. There are two major prison establishments in Sweden where those sentenced for driving under the influence are typically placed. They have a capacity of holding about 150 inmates altogether. These establishments have the security class 3, on a 1 to 3 scale where 3 is the most open establishment and where no concrete measures are taken to stop escapes. In fact, the facilities are only locked at night.

When it comes to the administrative procedure of the BAC test, an individual has to perform a screening breathalyzer test when caught in a police control. The test signals if the BAC level is above 0.2. If the test result turnout to be positive, the individual is required to take two additional Breathalyzer tests. The proof material in court is the average of these two new tests, less a 0.15 reduction (to account for e.g. consumption of food or medicine with small amounts of alcohol). If the individual refuses to take the Breathalyzer test, she needs to go to the police station to take a blood test instead.[5] A blood test is also taken if there is reason to suspect an intake of illegal drugs.

Turning to our data sources, data on Breathalyzer tests and data on blood tests come from the National Forensics Centre (NFC) while the crime data originate from The Swedish National

[3] For levels just above 0.2, the driving license may not be withdrawn, but above 0.32, the driving license will be immediately confiscated and the Swedish Transport Agency will decide about future withdrawal.
[4] The highest BAC level 1.5 used to be a law until 1994. It is now a norm which the Swedish courts follow (NJA, 2002).
[5] This contrasts with the US system where refusal leads to a penalty as if the BAC level was high.

Council for Crime Prevention. The Breathalyzer and blood test data cover the universe of all individuals caught with a BAC level of above 0.2 between 2008 and 2012 in Sweden. Table 1 reveals that there about 46,000 observations in Breathalyzer data and about 19,000 observations in the blood data.[6] The average BAC levels are much higher in the blood data (0.86 vs 1.46), which is a consequence of individuals tending to refuse the Breathalyzer test the larger the alcohol intake.

The next step in the empirical analysis is to match the individuals in the Breathalyzer and the blood data with the data on sentencing. Only the most severe crime is rendering an official punishment in Sweden. We therefore need to restrict our analysis to individuals that are being sentenced with drunk driving as their most severe crime since otherwise it is impossible to link a specific punishment to a specific crime. In the sentencing data, there are 41,437 cases where drunk driving is the main crime and we can match 33,199 of these cases with the Breathalyzer and the blood data combined, which gives us a match rate of 80 percent. Finally, we only analyze working adults (age 18-65) to avoid other sentencing-based thresholds in the Swedish judiciary system. Thus, we therefore end up with 29,905 observations in our analysis. Table 1 provides summary statistics of the data sets.

Table 1. Description of the data sets

Data	***Observations***	***BAC-level***	***St.dev***	***Min***	***Max***
Breathalyzer data	45,630	0.86	0.54	0.2	3.246
Blood data	19,410	1.46	0.73	0.2	4.5
Estimating sample	29,905				

Notes. The Breathalyzer data has been scaled with a factor of two to be comparable with blood data.

Table 2 shows the crime data for the sample of 30,083 individuals. The crime data include the following four set of categories (i) traffic (ii) drugs (iii) violence and (iv) theft/burglary. Traffic is by far the most common type of crime committed. We will use the total number of crimes as our key dependent variable. The baseline characteristics are age at crime, gender, immigrants, income and unemployment. Only 14 percent are females and 21 percent are immigrants. Income is measured as annual income from labor before tax one year before the crime and is measured in 100 of SEK. Unemployment is annual unemployment benefits in 100 of SEK.

[6] 1786 observations were duplicates, i.e., the same person and date occurring at least twice in total and they are dropped. After dropping the duplicate combinations of offenders and date we are left with 65,040 observation.

Table 2. Summary statistics for crime data and baseline characteristics

Crime categories	***Mean***	***St.dev***	***Min***	***Max***
Traffic	0.202	0.653	0	14
Drugs	0.0554	0.402	0	13
Violent	0.0331	0.206	0	6
Theft/burglary	0.0442	0.355	0	18
Total number of crimes	0.335	1.129	0	28
Baseline characteristics				
Age at crime	43.58	13.42	18	65
Female	0.137	0.344	0	1
Immigrants	0.212	0.409	0	1
Income	1985	2010	0	44,089
Unemployment benefits	54.20	211.0	0	1,986

Notes. The number of individuals included in this data set is 29,905.

3. The empirical design

In this section, we describe our empirical approach. At a general level, we are interested in estimating the following population regression model

(1) $$Crime_i = \alpha + \beta P_i + u_i,$$

where $Crime_i$ is a measure of future criminal behavior of individual i and P_i is an indicator for being sentenced to some specific form of punishment (e.g., imprisonment, probation, fees) and u_i is an error term that include all other factors that affect crime. The parameter of interest is β, which measures the causal effect of the specific punishment relatively to some other punishment on recidivism. We would expect that $\beta<0$ if the specific punishment deters future crime in relation to the alternative punishment.

One of the empirical problem in estimating equation (1) is that the likelihood of being sentenced to a specific form of punishment for a certain individual i depends on unobservable characteristics of either the crime committed or the individuals past criminal record. Thus, without a plausible exogenous source of variation in punishment, P_i will be correlated with unobserved characteristics in the error term, i.e., Cov(u,P)≠0. Another problem is that counterfactual punishment (P_i=0) is typically not well-defined since a judge has many types of punishments at her disposal as noted previously. However, thanks to our particular RD set-up we can overcome both of these problems. Specifically, our RD design allows us to compare two types of punishment, namely probation versus imprisonment, which include electronic monitoring. Since our RD design is fuzzy it can also be expressed by the following two equations:

(2) $P_i = a + bZ_i + f(BAC_i) + v_i$

(3) $Crime_i = c + dZ_i + g(BAC_i) + w_i$

where Z_i is an indicator for being above the BAC threshold. The casual effect can be measured at the threshold as the ratio of estimate of b in equation (2) and d in equation (3). This is equivalent to a Wald estimator, where equation (1) is the first-stage and equation (2) is the reduced form relationship, as first discussed by Hahn et al. (2001).

There are a number of different ways that one can implement and present the results from a RD design (e.g., Imbens and Lemieux, 2008, Lee and Lemieux, 2010 and Calonico et al., 2016). In this paper we will go through the following steps in our RD analysis: (i) testing whether the pre-treatment characteristics (age, income, unemployment, gender and age at migration) including the baseline outcomes (past criminal behavior) are balanced, (ii) testing for sorting using the McCrary density test (2012), (iii) showing results where we pool the observations across the two BAC thresholds (1.0 and 1.5) in order to significantly increase statistical power, (iv) showing the results from a large number of bandwidths including the optimal one. Specifically, we will present bandwidths starting at 0.02 percentage point, which is an extremely small neighborhood around the threshold (BAC of 0.02 percentage points is equivalent to a small fraction of a unit of alcohol), (v) showing results from both local linear and quadratic polynomial specifications, (vi) showing graphical evidence.

4. Results

In this section, we report results from the RD designs. We start by showing the test of balance for the pooled sample. Table 3 shows these results for all baseline (pretreatment) characteristics including the baseline outcomes. We use a local linear regression estimator with a uniform kernel. A large number of bandwidths are displayed in columns 1 to 7 starting with 0.02 and ending with 0.26. Colum 8 shows the results from the optimal bandwidth using the RD estimator developed by Calonico et al. (2016). Figure 1 shows the corresponding graphical evidence. Table 3 reveals that all baseline characteristics are balanced around the BAC threshold.

A test of discontinuity in the density of the forcing variable using the McCrary (2008) test does not indicate any sorting since the estimated effect is 0.055 with a standard error of 0.050. Figure 2 shows the graphical evidence.

Turning to the first-stage estimates, Table 4 displays these results in exactly the same way as in Table 3. The only differences are that we also show results from a local quadratic regression estimator and present results with (Panel B) and without (Panel A) the baseline covariates. Figure 2 shows the graph of the first-stage relationship. Table 8 shows that the first-stage estimate is about 10 percentage points, i.e., there is a jump in the probability of imprisonment of 10 percentage points at the BAC threshold. Reassuringly, the estimated size of the effect is very robust to choice of bandwidth, polynomial specifications as well as inclusion of baseline covariates.

Table 5 display the reduced form relationship for the total number of crimes committed in the future after the punishment period. We see that estimated effect is about −0.10 which implies that individuals commit 30% fewer crimes (relative a baseline of 0.335) if they have been sentenced to imprisonment rather than probation. Again, the estimated effect is robust across most of the specifications.

For completeness, Table 6 shows the instrumental variable results, i.e., the Wald estimator, which is the ratio of the reduced form effect and the first-stage estimate.

Table 7 shows the reduced form results for the four different crimes categories (i) traffic (ii) drugs (iii) violence and (iv) theft/burglary. Here we only show the results from the local linear regression estimator including baseline covariates. Table 7 reveals that all crime categories with the exception of theft (burglary) are reduced, although some of the estimates are somewhat imprecise.

5. Conclusions

Models of specific deterrence hold that the negative experience of imprisonment will tend to reduce the willingness to commit crime in the future, and the harsher prison conditions the larger will this effect be. However, the earlier literature using exogenous variation in prison standards has shown the opposite. Harsher conditions, if anything, reduces the propensity to commit crime in the future indicating that negative peer effects and stigmatization dominates the deterrent effect if it exists. This calls for the question if non-custodial sanctions would reduce future crime even further. We therefore compare sentences for drunk driving in minimum-security prisons or electronic monitoring with non-custodial sanctions. Interestingly we find that, at least for drunk driving, there is a lower bound for this result. The optimal punishment in terms of reducing future crime rates as much as possible seems to be minimum-security prisons. We note finally that more studies focusing on other types of crime and using exogenous variation in the penalty would be very useful for policy makers.

References

Abrams, David S. (2011),"Building Criminal Capital vs Specific Deterrence: The Effect of Incarceration Length on Recidivism", Mimeo, University of Pennsylvania.

Aizer, Anna and Doyle, John (2015), "Juvenile Incarceration, Human Capital and Future Crime: Evidence from Randomly Assigned Judges, *Quarterly Journal of Economics*, 130: 759-803.

Barbarino, Alessandro, Mastrobuoni, Giovanni (2014), "The Incapacitation Effect of Incarceration: Evidence from Several Italian Collective Pardons", *American Economic Journal: Economic Policy*, 6: 1-37.

Berube, Danton A., and Green Donald P. (2007), "The Effects of Sentencing on Recidivism: Results from a Natural Experiment", 2nd Annual Conference on Empirical Legal Studies Paper.

Buonanno, Paulo and Raphael, Steven (2013), "Incarceration and Incapacitation: Evidence from the 2006 Italian Collective Pardon", *American Economic Review*, 103: 2437-2465.

Calonico, Cattaneo, Farrell and Titiunik (2016), "Rdrobust: Software for Regression Discontinuity Designs," mimeo, University of Michigan.

Chen, Keith M. and Shapiro, Jesse (2007), "Do Harsher Prison Conditions Reduce Recidivism? A Discontinuity-based Approach", *American Law and Economics Review*, 9: 1-29.

Drago, Franceso, Galbiati, Roberto and Vertova, Pietro (2009), "The Deterrent Effects of Prison: Evidence from a Natural Experiment", *Journal of Political Economy*, 117: 257-280.

Green, Donald P. and Winik, Daniel (2010), "Using Random Judge Assignments to Estimate the Effects of Incarceration and Probation on Recidivism Among Drug Offenders", *Criminology*, 48: 357-387.

Hansen, Benjamin (2015), "Punishment and Deterrence: Evidence from Drunk Driving", *American Economic Review*, 105: 1581-1617.

Helland, Eric and Tabarrok, Alexander (2007), "Does Three Strikes Deter? A Non-Parametric Estimation", *Journal of Human Resources*, 22: 309-330.

Hinnerich Tyrefors, Björn, Palme, Mårten and Priks, Mikael (2016), "Imprisonment and Crime: Evidence using Discounts at Sentencing", mimeo, Stockholm University.

Hjalmarsson, Randi (2009a), "Juvenile Jails: A Path to the Straight and Narrow or to the Hardened Criminality", *Journal of Law and Economics*, 2009, 52: 779-809.

Hjalmarsson, Randi (2009b), "Crime and Expected Punishment: Changes in Perceptions at the Age of Criminal Majority", *American Law and Economics Review*, 11: 209-248.

Imbens, Guido, and Lemieux, Thomas (2008), "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142: 615-35.

Katz, Lawrence, Levitt, Steven D. and Shustorovich Ellen (2003), “Prison Conditions, Capital Punishment, and Deterrence”, *American Law and Economics Review,* 5: 318-343.

Kessler, Daniel och Levitt Steven (1999), ”Using Sentence Enhancements to Distinguish between Deterrence and Incapacitation”, *Journal of Law and Economics*, 42: 343–363.

Kuziemko, Ilyana (2013), ”How Should Inmates be Released from Prison? An Assesment of Parole Versus Fixed Sentence Regimes”, *Quarterly Journal of Economics*, 178: 371-424.

Lee, David S., and Lemieux, Thomas (2010), “Regression Discontinuity Designs in Economics,” *Journal of Economic Literature,* 48: 281-355.

Lee, David S. and McCrary, Justin (2009), ”The Deterrent Effect of Prison: Dynamic Theory and Evidence”, forthcoming, *Advances in Econometrics*.

Mastrobuoni Giovanni and Terlizzese Danielle (2015), “Rehabilitating Rehabilitation: Prison Conditions and Recidivism”, Mimeo, University of Essex.

Maurin, Eric och Ouss, Aurélie (2009), “Collective Sentence Reduction and Recidivism”, Mimeo, Harvard University.

Mueller-Smith, Michael (2015), “The Criminal and Labor Market Impacts of Incarceration”, Mimeo, Columbia University.

NJA, 2002, p. 653, www.lagen.nu

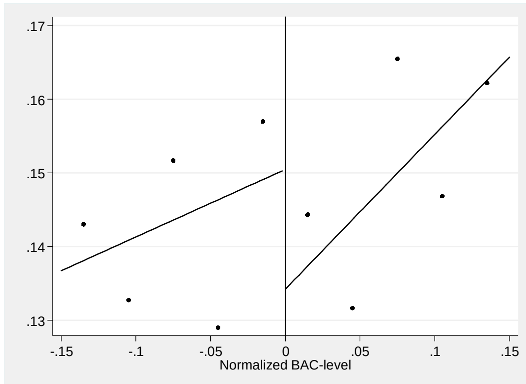
Owens, Emely G. (2009), ”More time, Less Crime? Estimating the Incapacitative Effect of Sentence Enhancements”, *Journal of Law and Economics,* 52: 551-579.

Raphael, Steven och Ludwig, Jens (2003), ”Prison Sentence Enhancements: The Case of Project Exile.” i Jens Ludwig and Philip J. Cook. (red.), *Evaluating Gun Policy: Effects on Crime and Violence*, Washington, DC: Brookings Institution Press.

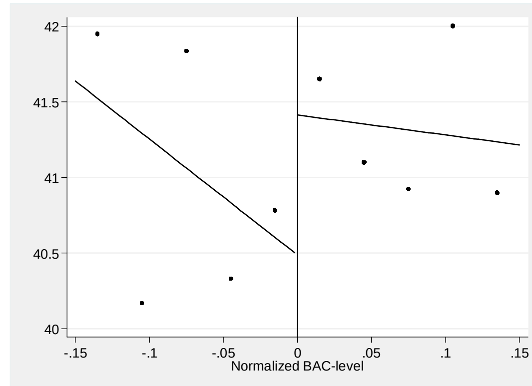
van der Werff, Cornelia (1979). Speciale Preventive, Den Haag: Staatsuitgeverij.

Figure 1: Graphs Baseline (pretreatment) characteristics

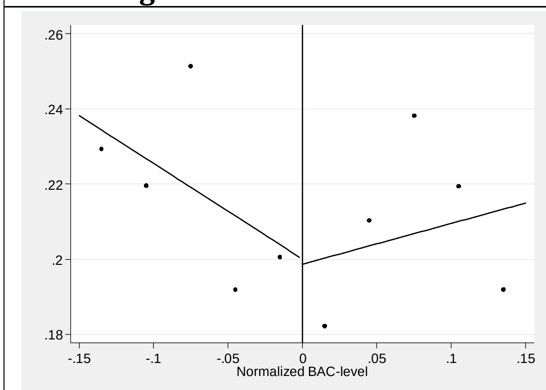
Share female



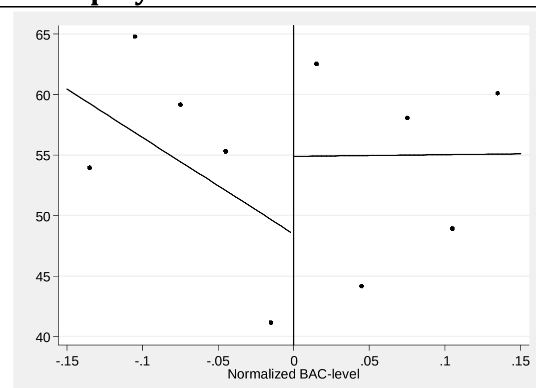
Age at crime



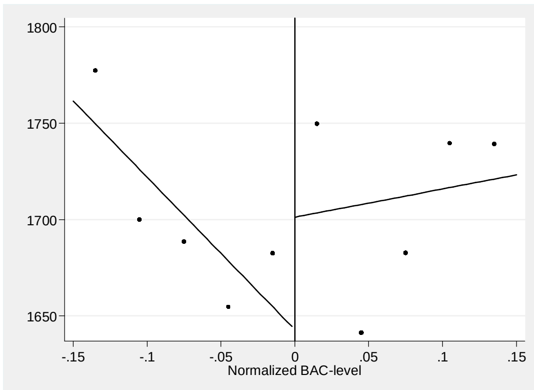
Share migrants



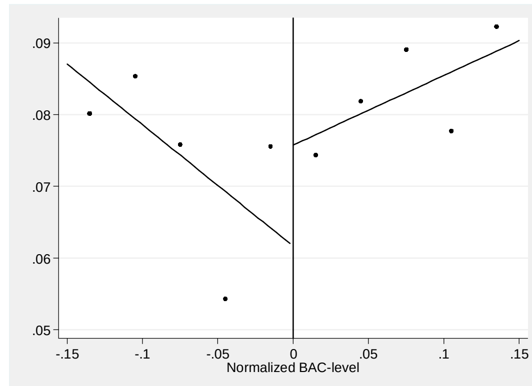
Unemployed



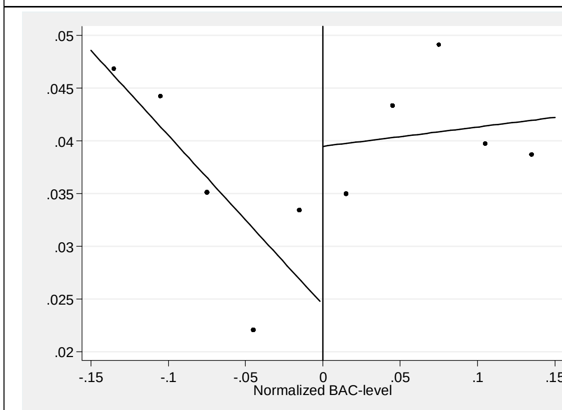
Labor income



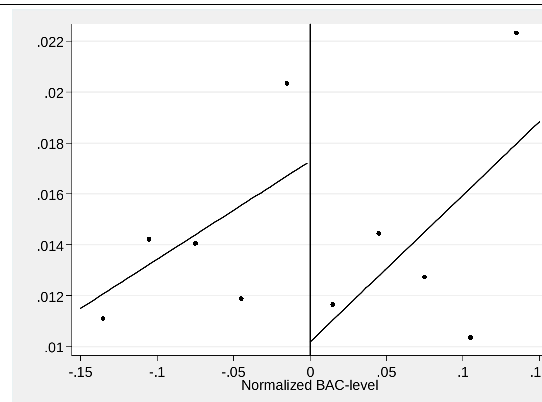
Total crimes within one year before the verdict



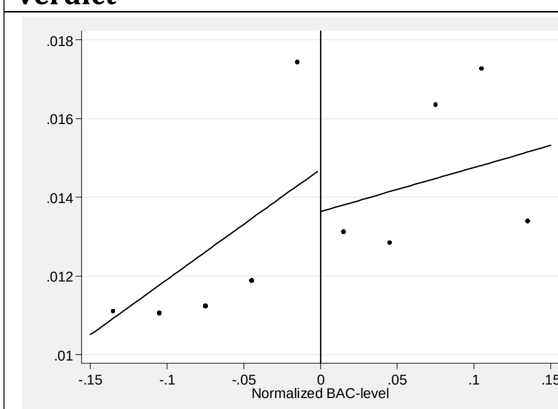
Traffic crimes within one year before verdict



Drug crimes one year before verdict



Violence crimes one year before verdict



Burglary/theft crime one year before verdict

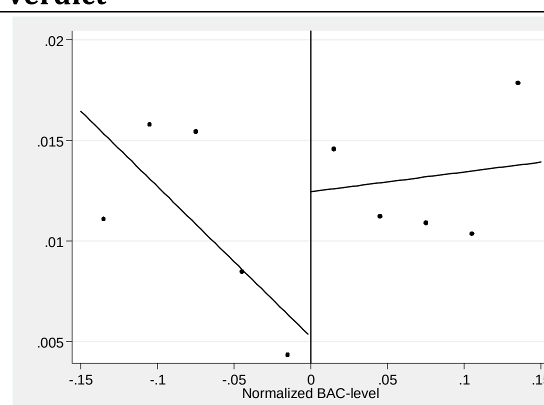
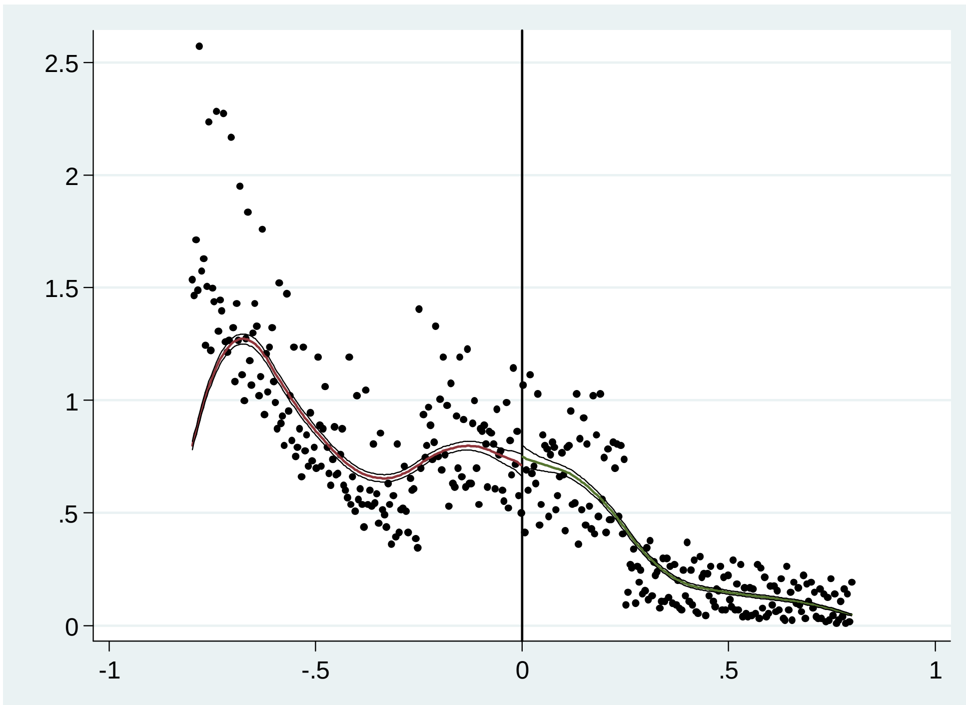
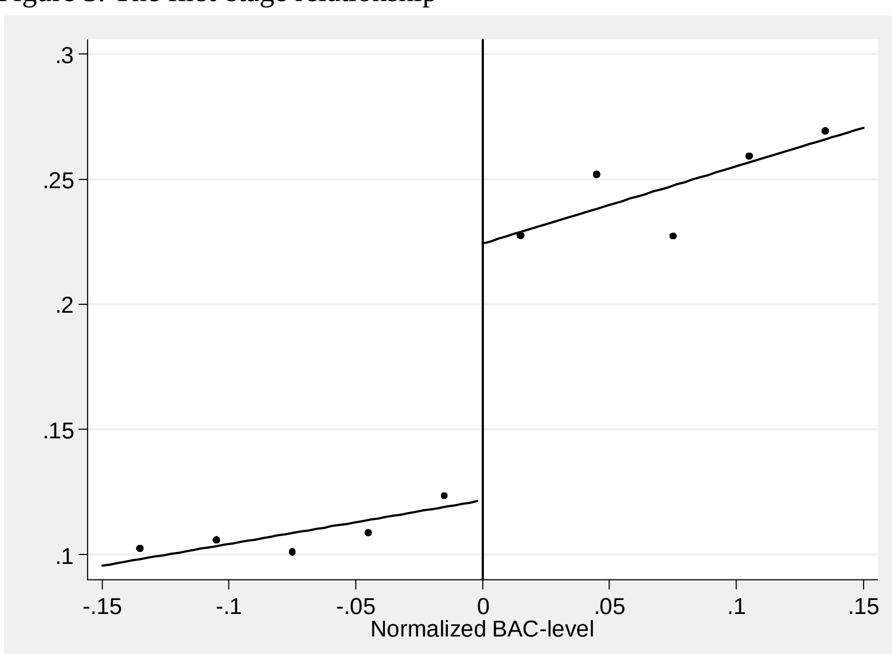


Figure 2. McCrary density plot.



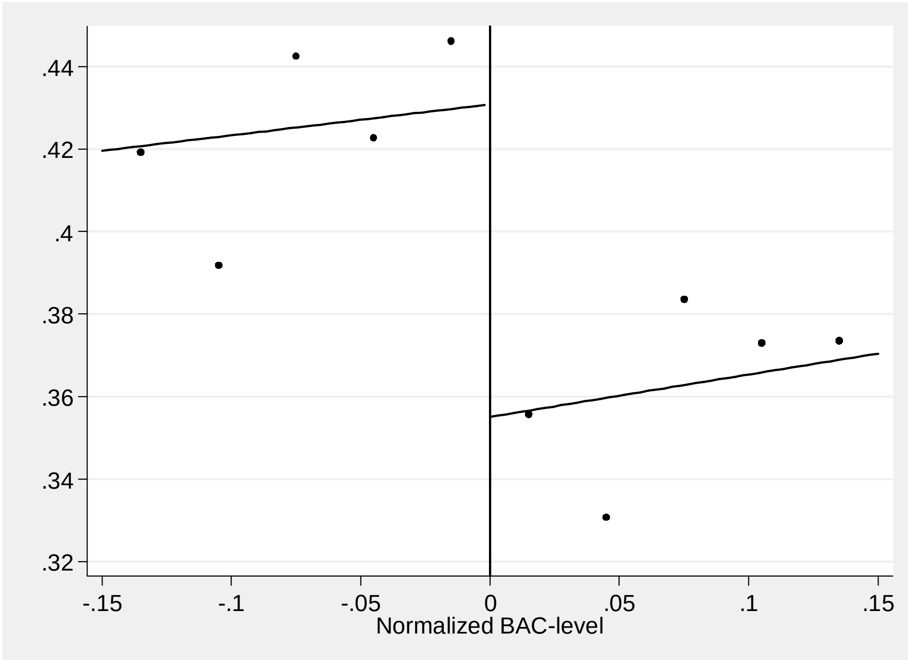
Notes: The estimated effect is 0.055 with a standard error of 0.050.

Figure 3. The first-stage relationship



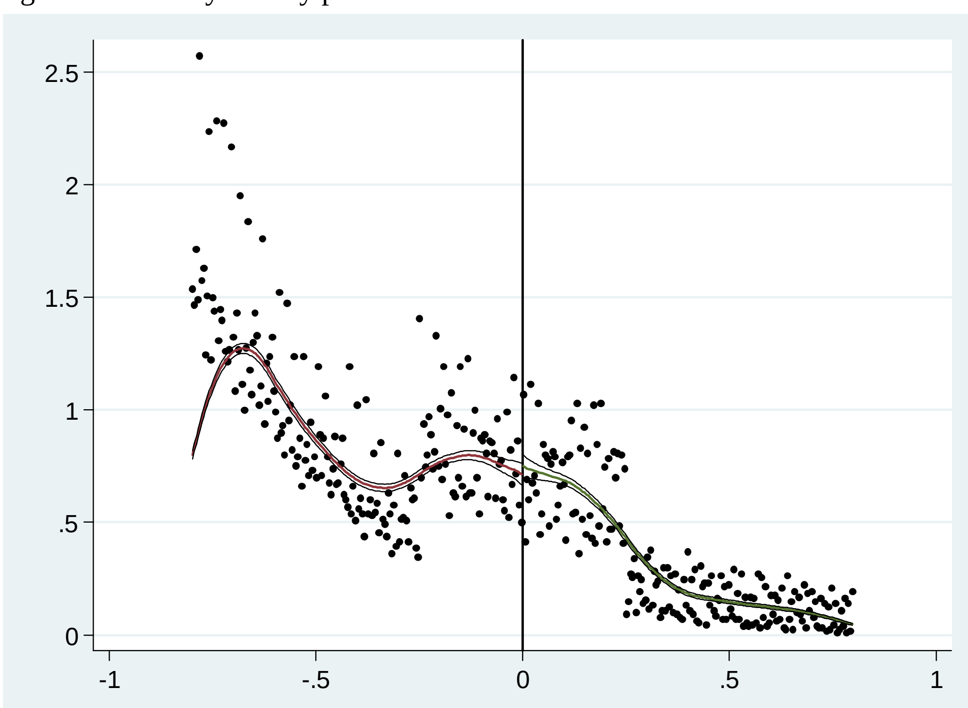
Notes: Bin size is 0.03. Bandwidth are 0.15 kernel(rec) bwidth(.15) degree(1) n(50).

Figure 4. The reduced form relationship



Notes: Bin size is 0.03. Bandwidth are 0.15. kernel(rec) bwidth(.15) degree(1) n(50)

Figure 5. McCrary density plot.



Notes: The estimated effect is 0.055 with a standard error of 0.050.

Table 3. Balancing tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidths	***0.02***	***0.06***	***0.10***	***0.14***	***0.18***	***0.22***	***0.26***	***Optimal***
				Age at crime				
BAC > threshold	-0.460	1.152	1.063	0.844	0.915	0.924	0.948*	0.924
	(2.003)	(1.121)	(0.868)	(0.721)	(0.639)	(0.579)	(0.534)	(0.606)
				Female				
BAC > threshold	-0.003	-0.026	-0.017	-0.012	-0.006	-0.013	0.002	-0.013
	(0.046)	(0.028)	(0.022)	(0.018)	(0.016)	(0.015)	(0.013)	(.015)
				Migrants				
BAC > threshold	-0.039	-0.036	-0.009	0.004	0.001	-0.000	-0.008	-0.002
	(0.054)	(0.031)	(0.025)	(0.020)	(0.018)	(0.017)	(0.015)	(0.017)
				Unemployed				
BAC > threshold	3.048	24.821	23.008*	8.255	2.393	-2.328	1.319	10.395
	(30.497)	(16.437)	(12.683)	(10.502)	(9.456)	(8.572)	(8.025)	(10.315)
				Labor income				
BAC > threshold	-20.731	74.504	36.924	37.799	62.848	75.315	24.669	54.85
	(233.566)	(137.796)	(108.306)	(90.436)	(83.076)	(73.597)	(67.239)	(69.465)
				Total number of crimes				
BAC > threshold	0.033	-0.013	0.008	0.002	0.008	0.018	0.009	0.011
	(0.047)	(0.027)	(0.022)	(0.017)	(0.015)	(0.015)	(0.014)	(0.015)
				Traffic crimes				
BAC > threshold	-0.007	-0.012	0.004	0.009	0.008	0.012	0.009	0.008
	(0.024)	(0.015)	(0.013)	(0.010)	(0.009)	(0.010)	(0.009)	(0.009)
				Narcotics crimes				
BAC > threshold	-0.000	-0.012	-0.005	-0.011	-0.007	-0.002	-0.001	-0.003
	(0.025)	(0.013)	(0.009)	(0.007)	(0.006)	(0.006)	(0.005)	(0.006)
				Violent crimes				
BAC > threshold	0.023	-0.000	-0.003	-0.002	-0.001	0.001	0.000	(.001)
	(0.017)	(0.009)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)
				Theft/burglary crimes				
BAC > threshold	0.016	0.010	0.013*	0.006	0.007	0.008*	0.000	0.005
	(0.013)	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)
N	917	2586	4225	6100	7770	9488	11154	

Notes: Each entry is a separate regression. In columns 1-7, we have used local linear regressions with a rectangular kernel. In Column 8, we have used the rdrobust as developed by Colonico et al (2016). Heteroscedasticity-robust standard errors are within parentheses. Coefficients significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 4. First-stage estimates

	Polynomial	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidths		***0.02***	***0.06***	***0.10***	***0.14***	***0.18***	***0.22***	***0.26***	***Optimal***
						Panel A: No Controls			
BAC > threshold	1	0.009	0.077***	0.108***	0.103***	0.101***	0.095***	0.083***	0.096***
		(0.051)	(0.030)	(0.023)	(0.019)	(0.017)	(0.016)	(0.014)	(0.021)
	2	0.124	0.023	0.057	0.084***	0.094***	0.098***	0.117***	0.106***
		(0.078)	(0.046)	(0.035)	(0.029)	(0.026)	(0.023)	(0.022)	(0.022)
N		917	2586	4225	6100	7770	9488	11154	
						Panel B: With Pretreatment controls			
BAC > threshold	1	-0.009	0.071**	0.104***	0.103***	0.102***	0.098***	0.084***	0.096***
		(0.054)	(0.030)	(0.023)	(0.019)	(0.017)	(0.016)	(0.014)	(0.021)
	2	0.083	0.006	0.055	0.083***	0.094***	0.097***	0.118***	0.100***
		(0.084)	(0.047)	(0.036)	(0.029)	(0.026)	(0.023)	(0.022)	(0.022)
N		906	2559	4169	6019	7670	9368	11012	

Notes: The dependent variable is an indicator for being above the 1.0 or 1.5 BAC thresholds. Each entry is a separate regression. In columns 1-7, we have used local linear regressions with a rectangular kernel. In Column 8, we have used the rdrobust as developed by Colonico et al (2016). The optimal bandwidth in Column 8 are 0.143, 0.289, 0.143 and 0.271 respectively. Heteroscedasticity-robust standard errors are within parentheses. Coefficients significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 5. The reduced-form estimate.

	Polynomial	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidths		***0.02***	***0.06***	***0.10***	***0.14***	***0.18***	***0.22***	***0.26***	***Optimal***
					Panel A: No Controls				
BAC > threshold	1	-0.162	-0.115	-0.092	-0.103*	-0.088*	-0.108**	-0.129***	-0.106**
		(0.173)	(0.092)	(0.070)	(0.053)	(0.049)	(0.045)	(0.043)	.04519
	2	-0.269	-0.231	-0.195*	-0.126	-0.108	-0.080	-0.073	-0.090
		(0.273)	(0.155)	(0.109)	(0.088)	(0.075)	(0.066)	(0.060)	(0.065)
N		917	2586	4225	6100	7770	9488	11154	
					Panel B: With Pretreatment controls				
BAC > threshold	1	-0.257*	-0.107	-0.122**	-0.103**	-0.088**	-0.112***	-0.118***	-0.106**
		(0.147)	(0.081)	(0.059)	(0.047)	(0.044)	(0.041)	(0.039)	(0.045)
	2	-0.253	-0.266**	-0.188**	-0.162**	-0.117*	-0.075	-0.087	-0.092
		(0.257)	(0.131)	(0.092)	(0.075)	(0.064)	(0.058)	(0.053)	(0.059)
N		906	2559	4169	6019	7670	9368	11012	

Notes: The dependent variable is the total number of future crimes. Each entry is a separate regression. In columns 1-7, we have used local linear regressions with a rectangular kernel. In Column 8, we have used the rdrobust as developed by Colonico et al (2016). The optimal bandwidth in Column 8 are 0.253, 0.270, 0.253 and 0.248 respectively. Heteroscedasticity-robust standard errors are within parentheses. Coefficients significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 6. IV-estimates

	Polynomial	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidths		***0.02***	***0.06***	***0.10***	***0.14***	***0.18***	***0.22***	***0.26***	***Optimal***
					Panel A: No Controls				
BAC > threshold	1	-28.313	-1.376	-0.929	-1.017*	-0.859*	-1.086**	-1.469***	-1.099**
		(209.680)	(1.305)	(0.670)	(0.540)	(0.490)	(0.498)	(0.566)	(.503)
	2	-3.318	-12.885	-3.219	-1.634	-1.165	-0.869	-0.662	-0.898
		(4.113)	(33.348)	(2.538)	(1.160)	(0.826)	(0.688)	(0.520)	(0.686)
N		917	2586	4225	6100	7770	9488	11154	
					Panel B: With Pretreatment controls				
BAC > threshold	1	29.806	-1.503	-1.172*	-0.994**	-0.866*	-1.151**	-1.400***	-1.026**
		(187.647)	(1.316)	(0.619)	(0.496)	(0.453)	(0.461)	(0.515)	(0.463)
	2	-3.059	-42.744	-3.452	-1.947*	-1.246	-0.782	-0.735	-0.975
		(4.410)	(320.901)	(2.811)	(1.121)	(0.767)	(0.628)	(0.470)	(0.671)
N		906	2559	4169	6019	7670	9368	11012	

Notes: The dependent variable is the total number of future crimes. Each entry is a separate regression. In columns 1-7, we have used local linear regressions with a rectangular kernel. In Column 8, we have used the rdrobust as developed by Colonico et al (2016). The optimal bandwidths in Column 8 are 0.253, 0.270, 0.231 and 0.248 respectively. Heteroscedasticity-robust standard errors are within parentheses. Coefficients significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 7. The reduced-form estimates different crimes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidths	***0.02***	***0.06***	***0.10***	***0.14***	***0.18***	***0.22***	***0.26***	***Optimal***
			Total number of traffic related crime (drunk driving included)					
BAC > threshold	-0.065	-0.048	-0.068*	-0.052	-0.047	-0.055**	-0.046*	-0.048*
	(0.097)	(0.051)	(0.040)	(0.032)	(0.029)	(0.027)	(0.025)	(0.029)
			Total number of narcotics related crime					
BAC > threshold	-0.157**	-0.055	-0.038*	-0.033*	-0.034**	-0.037**	-0.045***	-0.044***
	(0.065)	(0.035)	(0.023)	(0.018)	(0.017)	(0.015)	(0.015)	(0.017)
			Total number of violence related crime					
BAC > threshold	-0.057*	-0.041**	-0.026*	-0.018	-0.010	-0.013	-0.020**	-0.017
	(0.032)	(0.017)	(0.014)	(0.012)	(0.010)	(0.009)	(0.009)	(0.010)
			Total number of burglary and theft related crime					
BAC > threshold	0.022	0.037*	0.010	-0.000	0.003	-0.007	-0.007	0.008
	(0.035)	(0.022)	(0.017)	(0.014)	(0.013)	(0.013)	(0.011)	(0.015)
N	906	2559	4169	6019	7670	9368	11012	

Notes: Each entry is a separate regression. In columns 1-7, we have used local linear regressions with a rectangular kernel. In Column 8, we have used the rdrobust as developed by Colonico et al (2016). The optimal bandwidths in Column 8 are 0.251, 0.262, 0.208 and 0.187 respectively. Heteroscedasticity-robust standard errors are within parentheses. Coefficients significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.