

# Effects of Incarceration on Recidivism and Labor Market Outcomes and Ban the Box

Incarceration, Labor Market, and Ban the Box

Hussain Hadah (he/him)

October 13, 2025



Tulane  
University

# Outline for Today

1. **Effect of Incarceration**
2. **Summary of Bhuller et al. (2020)**
3. **Summary of Eren, Mocan (2021)**
4. **Introduce Audit Studies**
5. **Ban the Box**



# This week

- We will start housing policy on Friday







# Next week

- Quiz 3 on Monday, October 20th
- Same instructions as before
- More housing!

## Readings

- Read all of Metcalf (2018) "*Sand Castles before the Tide? Affordable Housing in Expensive Cities*"
- Skim Gyourko (1990)
- Skim Clark and Freedman (2019)
- Skim Aaronson et. al. (2021)
- Skim Greenaway-McGrevy and Phillips (2023)
- Skim Song (2021)
- Skim Jarvis (2021)
- Hanlon and Heblich (2022) sections 4.4 and 4.5 only



# Measuring the effects of incarceration



# How Does Incarceration Affect Recidivism, and Labor Market Outcomes?

- The idea is that incarceration, i.e. putting people in jail, reduces the chances that they engage in criminal activity later
- Is this the case?
- First, let's summarize some possible mechanisms of how incarceration could affect recidivism and labor market outcomes

# How Does Incarceration Affect Recidivism, and Labor Market Outcomes?

- The idea is that incarceration, i.e. putting people in jail, reduces the chances that they engage in criminal activity later
- Is this the case?
- First, let's summarize some possible mechanisms of how incarceration could affect recidivism and labor market outcomes

What are some possible ways of how incarceration could affect recidivism and labor market outcomes?



# How Does Incarceration Affect Recidivism, and Labor Market Outcomes?

- Incarceration could **reduce** recidivism by:
  - Reducing the temptation to engage in criminal activity later, to avoid the costs of incarceration that were just experienced (i.e. a deterrent effect)
  - Increasing education or job training, to the extent that those programs are available in prison. This could increase job market opportunities which would decrease incarceration (think the “rational criminal” model)
  - Crime often occurs due to substance misuse or mental health issues. If incarceration reduces, rather than exacerbates, those issues then this could reduce recidivism and boost labor supply

# How Does Incarceration Affect Recidivism, and Labor Market Outcomes?

- Incarceration could **increase** recidivism by:
  - Causing a deterioration in human capital (i.e. skills, work experience) since the individual is generally not working. This needs to be contrasted with any education or training programs available in prison. The net effect could go either way
  - Those with criminal records face discrimination in the job market, which reduces their ability to get a job. This leads to recidivism for many
  - If incarceration increases substance misuse or mental health issues, that could reduce labor supply and increase recidivism

# How Does Incarceration Affect Recidivism, and Labor Market Outcomes?

- Determining what the net effect is of incarceration – does it increase or decrease recidivism/labor market outcomes – is tricky because incarceration is not randomly assigned
- Those who are sent to prison are different from those who are not
- A simple comparison of, say, labor market outcomes between those who faced incarceration and those who did not would give a biased estimate of the “effect” of incarceration on labor market outcomes
- E.g., those who faced incarceration had worse labor market outcomes anyways
- The ideal would be to compare two on-average identical groups of people: one group is put in prison and the other group is not
- This sort of randomized control trial is obviously not possible or ethical
- So, economists and social scientists look to “natural experiments”, or ways that there was quasi-random variation in incarceration

Bhuller et al. (2020)



# Abstract

- Understanding the impact of prison time on criminal behavior is complex due to limited data and differences between those incarcerated and those not
- The study focuses on Norway's criminal justice system, utilizing a comprehensive dataset on criminal behavior and labor market outcomes
- It leverages the varying strictness of judges (as an instrumental variable) to assess the effects of incarceration on reoffending and employment
- Findings show that imprisonment significantly decreases the likelihood of reoffending by 27 percentage points and reduces subsequent criminal charges by 10

# Abstract

- In contrast, ordinary least squares (OLS) analysis suggests a positive relationship between incarceration and future crime, highlighting the influence of selection bias
- The study clarifies that the observed high recidivism rates among ex-convicts are more about the pre-existing likelihood of committing crimes rather than the prison experience itself
- The preventive effect of incarceration is particularly notable among individuals not employed before imprisonment, indicating the importance of rehabilitation programs in reducing recidivism and improving employment outcomes
- These results challenge the 'nothing works' doctrine, suggesting that prison, when focused on rehabilitation, can have a preventive effect on criminal behavior



# Testing the Random Assignment of Judges

Table 1. Testing for Random Assignment of Criminal Cases to Judges.

|                                 | Dependent Variables: |          |                  |          | Explanatory Variables: |          |
|---------------------------------|----------------------|----------|------------------|----------|------------------------|----------|
|                                 | Pr(Incarcerated)     |          | Judge Stringency |          | (5)                    | (6)      |
|                                 | (1)                  | (2)      | (3)              | (4)      |                        |          |
|                                 | Coefficient          | Standard | Coefficient      | Standard |                        |          |
|                                 | Estimate             | Error    | Estimate         | Error    | Mean                   | Standard |
| Demographics and Type of Crime: |                      |          |                  |          |                        |          |
| Age                             | 0.0049***            | (0.0004) | -0.0000          | (0.0000) | 32.64                  | (11.35)  |
| Female                          | -0.0651***           | (0.0074) | -0.0011          | (0.0007) | 0.106                  | (0.308)  |
| Foreign born                    | 0.0084               | (0.0064) | 0.0007           | (0.0007) | 0.135                  | (0.342)  |
| Married, year t-1               | -0.0442***           | (0.0119) | -0.0018          | (0.0012) | 0.111                  | (0.514)  |
| Number of children, year t-1    | -0.0029              | (0.0033) | 0.0002           | (0.0004) | 0.783                  | (1.244)  |
| High school degree, year t-1    | -0.0675***           | (0.0134) | -0.0014          | (0.0015) | 0.172                  | (0.377)  |
| Some college, year t-1          | 0.0060               | (0.0084) | 0.0003           | (0.0009) | 0.046                  | (0.209)  |
| Violent crime                   | 0.0795***            | (0.0087) | 0.0014           | (0.0011) | 0.256                  | (0.437)  |
| Property crime                  | -0.0054              | (0.0115) | 0.0012           | (0.0012) | 0.139                  | (0.346)  |
| Economic crime                  | -0.0477***           | (0.0117) | 0.0018           | (0.0015) | 0.113                  | (0.516)  |
| Drug related                    | -0.0385***           | (0.0115) | 0.0000           | (0.0013) | 0.119                  | (0.324)  |
| Drunk driving                   | 0.0667***            | (0.0132) | 0.0001           | (0.0014) | 0.071                  | (0.257)  |
| Other traffic                   | -0.0457***           | (0.0127) | 0.0003           | (0.0012) | 0.087                  | (0.281)  |
| Missing Xs                      | -0.4905***           | (0.1415) | -0.0102          | (0.0152) | 0.030                  | (0.170)  |
| Past Work and Criminal History: |                      |          |                  |          |                        |          |
| Employed, year t-1              | 0.0104               | (0.0083) | 0.0001           | (0.0008) | 0.352                  | (0.478)  |
| Ever Employed, years t-2 to t-5 | -0.0055              | (0.0085) | 0.0001           | (0.0009) | 0.470                  | (0.499)  |
| Charged, year t-1               | 0.0932***            | (0.0074) | 0.0004           | (0.0008) | 0.459                  | (0.498)  |
| Ever Charged, years t-2 to t-5  | 0.0925***            | (0.0078) | -0.0006          | (0.0009) | 0.627                  | (0.483)  |
| F-statistic for joint test      | 60.61                |          | .577             |          |                        |          |
| [p-value]                       | [.000]               |          | [.917]           |          |                        |          |
| Number of cases                 | 33,509               |          |                  |          | 33,509                 |          |

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009. All estimations include controls for court x court entry year FEs. Reported F-statistic refers to a joint test of the null hypothesis for all variables. The omitted category for education is "Less than high school, year t-1" and the omitted category for type of crime is "Other crimes". Standard errors are two-way clustered at judge and defendant level. \*\*p<0.1, \*\*\*p<0.05, \*\*\*\*p<0.01.

Columns (1) and (2) show how demographics, type of crime, and past work and criminal history affect the probability that you become incarcerated

Not surprisingly, those who are incarcerated and different from those who are not

E.g., older, more likely to be male, single, without education, more likely to have done violent crime, and they have more criminal history

# Testing the Random Assignment of Judges

Table 1. Testing for Random Assignment of Criminal Cases to Judges.

|                                 | Dependent Variables: |          |                  |          | Explanatory Variables: |          |
|---------------------------------|----------------------|----------|------------------|----------|------------------------|----------|
|                                 | Pr(Incarcerated)     |          | Judge Stringency |          | (5)                    | (6)      |
|                                 | (1)                  | (2)      | (3)              | (4)      |                        |          |
|                                 | Coefficient          | Standard | Coefficient      | Standard |                        |          |
|                                 | Estimate             | Error    | Estimate         | Error    | Mean                   | Standard |
| Demographics and Type of Crime: |                      |          |                  |          |                        |          |
| Age                             | 0.0049***            | (0.0004) | -0.0000          | (0.0000) | 32.64                  | (11.35)  |
| Female                          | -0.0651***           | (0.0074) | -0.0011          | (0.0007) | 0.106                  | (0.208)  |
| Foreign born                    | 0.0084               | (0.0064) | 0.0007           | (0.0007) | 0.135                  | (0.342)  |
| Married, year t-1               | -0.0442***           | (0.0119) | -0.0018          | (0.0012) | 0.111                  | (0.514)  |
| Number of children, year t-1    | -0.0029              | (0.0033) | 0.0002           | (0.0004) | 0.783                  | (1.244)  |
| High school degree, year t-1    | -0.0675***           | (0.0134) | -0.0014          | (0.0015) | 0.172                  | (0.377)  |
| Some college, year t-1          | 0.0060               | (0.0084) | 0.0003           | (0.0009) | 0.046                  | (0.209)  |
| Violent crime                   | 0.0795***            | (0.0087) | 0.0014           | (0.0011) | 0.256                  | (0.437)  |
| Property crime                  | -0.0054              | (0.0115) | 0.0012           | (0.0012) | 0.139                  | (0.346)  |
| Economic crime                  | -0.0477***           | (0.0117) | 0.0018           | (0.0015) | 0.113                  | (0.516)  |
| Drug related                    | -0.0385***           | (0.0115) | 0.0000           | (0.0013) | 0.119                  | (0.324)  |
| Drunk driving                   | 0.0667***            | (0.0132) | 0.0001           | (0.0014) | 0.071                  | (0.257)  |
| Other traffic                   | -0.0457***           | (0.0127) | 0.0003           | (0.0012) | 0.087                  | (0.281)  |
| Missing Xs                      | -0.4905***           | (0.1415) | -0.0102          | (0.0152) | 0.030                  | (0.170)  |
| Past Work and Criminal History: |                      |          |                  |          |                        |          |
| Employed, year t-1              | 0.0104               | (0.0083) | 0.0001           | (0.0008) | 0.352                  | (0.478)  |
| Ever Employed, years t-2 to t-5 | -0.0055              | (0.0085) | 0.0001           | (0.0009) | 0.470                  | (0.499)  |
| Charged, year t-1               | 0.0932***            | (0.0074) | 0.0004           | (0.0008) | 0.459                  | (0.498)  |
| Ever Charged, years t-2 to t-5  | 0.0925***            | (0.0078) | -0.0006          | (0.0009) | 0.627                  | (0.483)  |
| F-statistic for joint test      | 60.61                |          | .577             |          |                        |          |
| [p-value]                       | [.000]               |          | [.917]           |          |                        |          |
| Number of cases                 | 33,509               |          |                  |          | 33,509                 |          |

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009. All estimations include controls for court x court entry year FEs. Reported F-statistic refers to a joint test of the null hypothesis for all variables. The omitted category for education is "Less than high school, year t-1" and the omitted category for type of crime is "Other crimes". Standard errors are two-way clustered at judge and defendant level. \*\*p<0.1, \*\*\*p<0.05, \*\*\*\*p<0.01.

Columns (3) and (4) show how these factors related to judge stringency

What we want to see if judge stringency is as-good-as-random is that there are few statistically significant relationships between judge stringency and demographic, crime, or work and criminal history factors

There is no association (nothing is statistically significant), so this strongly suggests that the assignment is as-good-as-random

# Summary Statistics

Table 2. Descriptive Statistics.

|                                             | (1)    | (2)                |
|---------------------------------------------|--------|--------------------|
|                                             | Mean   | Standard Deviation |
| <b><u>A. Defendant Characteristics:</u></b> |        |                    |
| Demographics:                               |        |                    |
| Age                                         | 33.07  | (11.78)            |
| Female                                      | 0.119  | (0.324)            |
| Foreign born                                | 0.148  | (0.355)            |
| Married, year t-1                           | 0.128  | (0.334)            |
| Number of children, year t-1                | 0.822  | (1.284)            |
| Some college, year t-1                      | 0.056  | (0.229)            |
| High school degree, year t-1                | 0.186  | (0.389)            |
| Less than high school, year t-1             | 0.758  | (0.428)            |
| Missing Xs                                  | 0.034  | (0.181)            |
| Past Work and Criminal History:             |        |                    |
| Employed, year t-1                          | 0.393  | (0.488)            |
| Ever Employed, years t-2 to t-5             | 0.505  | (0.499)            |
| Charged, year t-1                           | 0.378  | (0.485)            |
| Ever Charged, years t-2 to t-5              | 0.572  | (0.495)            |
| Number of defendants                        | 23,345 |                    |
| <b><u>B. Type of Crime:</u></b>             |        |                    |
| Violent crime                               | 0.256  | (0.437)            |
| Property crime                              | 0.139  | (0.346)            |
| Economic crime                              | 0.113  | (0.316)            |
| Drug related                                | 0.119  | (0.324)            |
| Drunk driving                               | 0.071  | (0.257)            |
| Other traffic                               | 0.087  | (0.281)            |
| Other crimes                                | 0.215  | (0.419)            |
| Number of cases                             | 33,509 |                    |

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009 with 23,345 unique defendants.

This descriptive statistics table tells you about the general make-up of their sample of individuals who are either assigned to incarceration or not

Defendants are more likely to be male, unmarried, have children, have low levels of education, have criminal history, and have low levels of employment history

# First Stage: Testing the Stringency of Judges

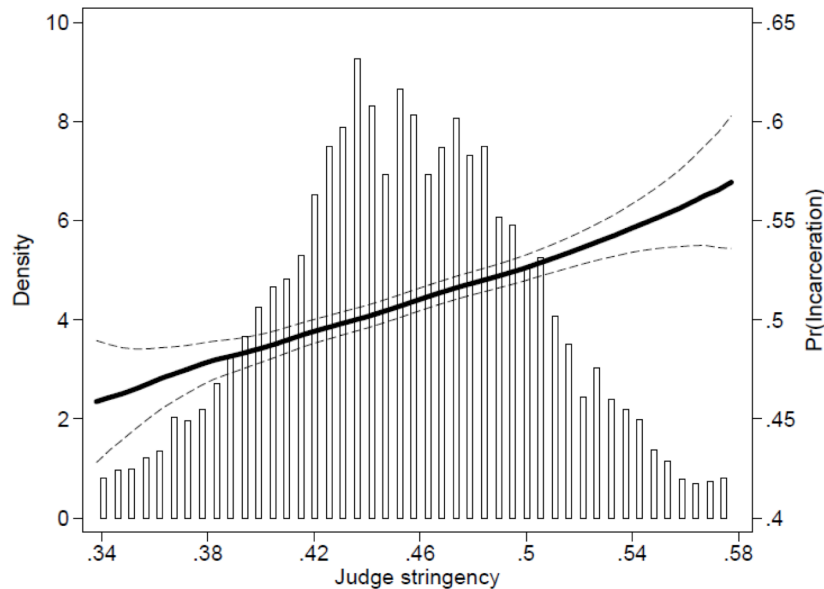


Figure 3. First Stage Graph of Incarceration on Judge Stringency.

Yes, as stringency increases, the probability of incarceration increases

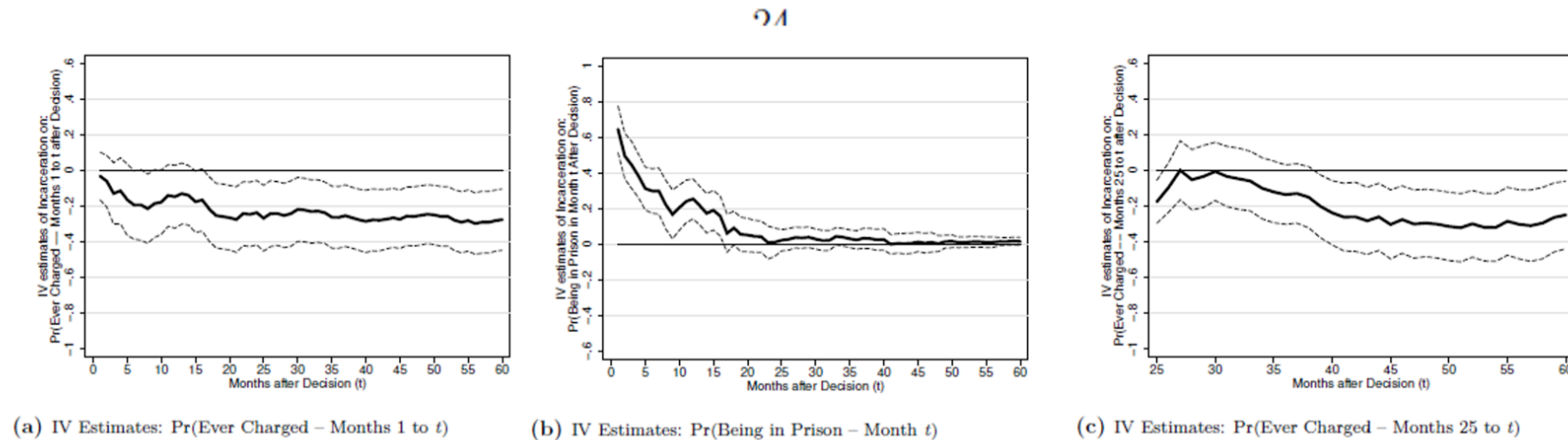
Range of probabilities is about 46% (most lenient) to 57% (most strict)

We can think of  $(57\% - 46\% = 11\%) / 57\% =$  about 19% of incarcerations occur due to judges being pickier than the more lenient judge

# The Effect of Incarceration on Recidivism

**Figure 4.** The Effect of Incarceration on Recidivism and Probability of Being in Prison.

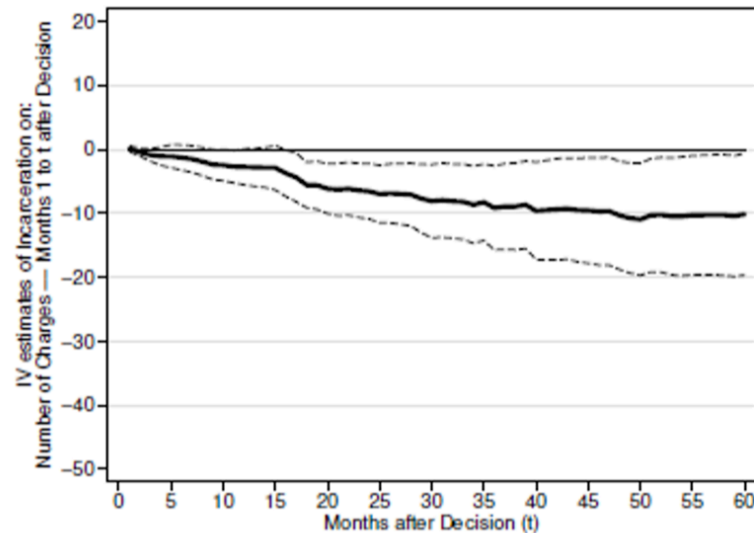
*Note:* Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Panel (b) plots prison probabilities related only to the original sentence. Dashed lines show 90% confidence intervals.



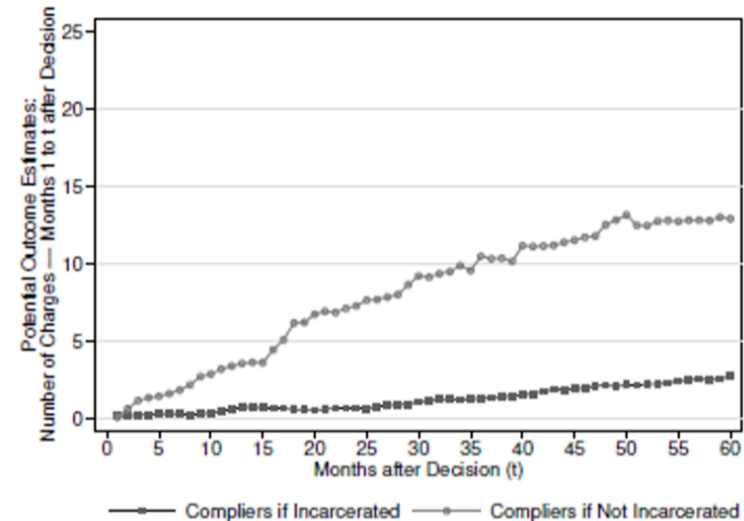
# The Effect of Incarceration on Recidivism

**Figure 5.** The Effect of Incarceration on Number of Charges.

*Note:* Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Dashed lines show 90% confidence intervals.



(a) IV Estimates: No. of Charges – Months 1 to  $t$



(b) Potential Outcomes: No. of Charges in Months 1 to  $t$

# The Effect of Incarceration on Recidivism

Table 5. The Effects of Incarceration on Recidivism.

| Dependent Variable:               | Pr(Ever Charged)                     |                                       |                                      | Number of Charges                    |
|-----------------------------------|--------------------------------------|---------------------------------------|--------------------------------------|--------------------------------------|
|                                   | Months 1-24<br>after Decision<br>(1) | Months 25-60<br>after Decision<br>(2) | Months 1-60<br>after Decision<br>(3) | Months 1-60<br>after Decision<br>(4) |
| OLS: Incarcerated                 | 0.130***                             | 0.113***                              | 0.113***                             | 5.092***                             |
| No controls                       | (0.007)                              | (0.007)                               | (0.006)                              | (0.315)                              |
| OLS: Incarcerated                 | 0.126***                             | 0.108***                              | 0.103***                             | 5.187***                             |
| Demographics & Type of Crime      | (0.007)                              | (0.007)                               | (0.006)                              | (0.303)                              |
| OLS: Incarcerated                 | 0.087***                             | 0.068***                              | 0.066***                             | 4.128***                             |
| All controls                      | (0.006)                              | (0.007)                               | (0.006)                              | (0.291)                              |
| OLS: Incarcerated                 | 0.083***                             | 0.065***                              | 0.066***                             | 3.735***                             |
| Complier Re-weighted              | (0.007)                              | (0.007)                               | (0.006)                              | (0.273)                              |
| RF: Judge Stringency              | -0.103**                             | -0.115**                              | -0.127***                            | -4.729*                              |
| All controls                      | (0.049)                              | (0.049)                               | (0.045)                              | (2.519)                              |
| IV: Incarcerated                  | -0.223*                              | -0.248**                              | -0.274***                            | -10.176*                             |
| All controls                      | (0.115)                              | (0.115)                               | (0.104)                              | (5.759)                              |
| Dependent mean                    | 0.57                                 | 0.56                                  | 0.70                                 | 9.91                                 |
| Complier mean if not incarcerated | 0.56                                 | 0.58                                  | 0.73                                 | 12.90                                |
| Number of cases                   | 31,287                               |                                       |                                      |                                      |

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1. RF and IV in addition also control for court x court entry year FEs. OLS standard errors are clustered at the defendant level, while RF and IV standard errors are two-way clustered at judge and defendant level. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

OLS estimates = Ordinary Least Squares. These are naïve estimates that just compare those incarcerated to those who are not

These show a positive association between incarceration and recidivism

Adding control variables cuts the effect in half, but even trying to control for observable factors to make those incarcerated = to those not incarcerated is an imperfect exercise

# The Effect of Incarceration on Recidivism

Table 5. The Effects of Incarceration on Recidivism.

| Dependent Variable:               | Pr(Ever Charged) |                |                | Number of      |
|-----------------------------------|------------------|----------------|----------------|----------------|
|                                   | Months 1-24      | Months 25-60   | Months 1-60    | Charges        |
|                                   | after Decision   | after Decision | after Decision | Months 1-60    |
|                                   | (1)              | (2)            | (3)            | after Decision |
| OLS: Incarcerated                 | 0.130***         | 0.113***       | 0.113***       | 5.092***       |
| No controls                       | (0.007)          | (0.007)        | (0.006)        | (0.315)        |
| OLS: Incarcerated                 | 0.126***         | 0.108***       | 0.103***       | 5.187***       |
| Demographics & Type of Crime      | (0.007)          | (0.007)        | (0.006)        | (0.303)        |
| OLS: Incarcerated                 | 0.087***         | 0.068***       | 0.066***       | 4.128***       |
| All controls                      | (0.006)          | (0.007)        | (0.006)        | (0.291)        |
| OLS: Incarcerated                 | 0.083***         | 0.065***       | 0.066***       | 3.735***       |
| Complier Re-weighted              | (0.007)          | (0.007)        | (0.006)        | (0.273)        |
| RF: Judge Stringency              | -0.103**         | -0.115**       | -0.127***      | -4.729*        |
| All controls                      | (0.049)          | (0.049)        | (0.045)        | (2.519)        |
| IV: Incarcerated                  | -0.223*          | -0.248**       | -0.274***      | -10.176*       |
| All controls                      | (0.115)          | (0.115)        | (0.104)        | (5.759)        |
| Dependent mean                    | 0.57             | 0.56           | 0.70           | 9.91           |
| Complier mean if not incarcerated | 0.56             | 0.58           | 0.73           | 12.90          |
| Number of cases                   | 31,287           |                |                |                |

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1. RF and IV in addition also control for court x court entry year FEs. OLS standard errors are clustered at the defendant level, while RF and IV standard errors are two-way clustered at judge and defendant level. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

RF/IV estimates = estimates that use judge fixed effects to compare on-average identical defendants who happened to be randomly assigned lenient vs strict judges

These estimates show the opposite effect – incarceration reduces recidivism

The decrease in recidivism persists even up to 60 months later



# The Effect of Incarceration on Recidivism

Table 5. The Effects of Incarceration on Recidivism.

| Dependent Variable:               | Pr(Ever Charged)              |                                |                               | Number of Charges             |
|-----------------------------------|-------------------------------|--------------------------------|-------------------------------|-------------------------------|
|                                   | Months 1-24<br>after Decision | Months 25-60<br>after Decision | Months 1-60<br>after Decision | Months 1-60<br>after Decision |
|                                   | (1)                           | (2)                            | (3)                           | (4)                           |
| OLS: Incarcerated                 | 0.130***                      | 0.113***                       | 0.113***                      | 5.092***                      |
| No controls                       | (0.007)                       | (0.007)                        | (0.006)                       | (0.315)                       |
| OLS: Incarcerated                 | 0.126***                      | 0.108***                       | 0.103***                      | 5.187***                      |
| Demographics & Type of Crime      | (0.007)                       | (0.007)                        | (0.006)                       | (0.303)                       |
| OLS: Incarcerated                 | 0.087***                      | 0.068***                       | 0.066***                      | 4.128***                      |
| All controls                      | (0.006)                       | (0.007)                        | (0.006)                       | (0.291)                       |
| OLS: Incarcerated                 | 0.083***                      | 0.065***                       | 0.066***                      | 3.735***                      |
| Complier Re-weighted              | (0.007)                       | (0.007)                        | (0.006)                       | (0.273)                       |
| RF: Judge Stringency              | -0.103**                      | -0.115**                       | -0.127***                     | -4.729*                       |
| All controls                      | (0.049)                       | (0.049)                        | (0.045)                       | (2.519)                       |
| IV: Incarcerated                  | -0.223*                       | -0.248**                       | -0.274***                     | -10.176*                      |
| All controls                      | (0.115)                       | (0.115)                        | (0.104)                       | (5.759)                       |
| Dependent mean                    | 0.57                          | 0.56                           | 0.70                          | 9.91                          |
| Complier mean if not incarcerated | 0.56                          | 0.58                           | 0.73                          | 12.90                         |
| Number of cases                   | 31,287                        |                                |                               |                               |

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1. RF and IV in addition also control for court x court entry year FEs. OLS standard errors are clustered at the defendant level, while RF and IV standard errors are two-way clustered at judge and defendant level. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

The decrease in recidivism is quite large

The IV estimate for months 1-60 is -0.274

The dependent mean (the baseline mean) “ever charged” probability for those not incarcerated is 0.70

Therefore, we can think of incarceration decreasing this probability from 0.70 (70%) to 0.70-0.274 = 0.426 (42.6%)

The incarceration rate decreases by 27.4 percentage points, or decreases by about 30.1% compared to the baseline rate of 70%

# The Effect of Incarceration on Recidivism

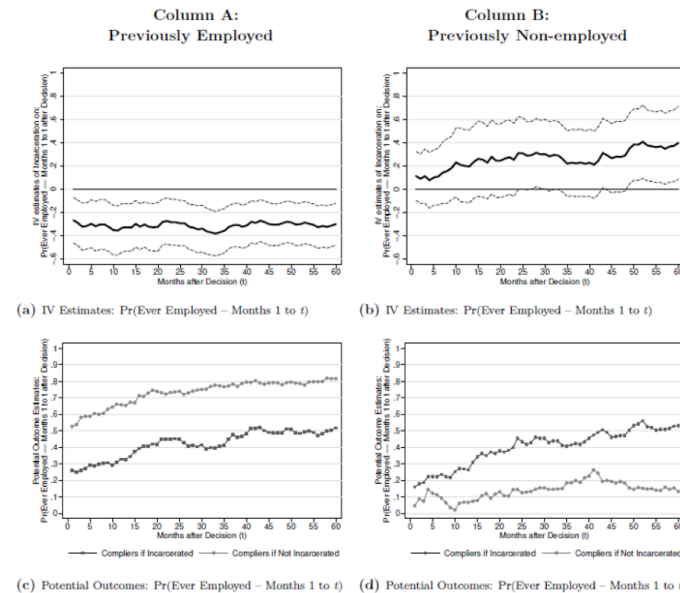


Figure 8. The Effect of Incarceration on Future Employment by Previous Labor Market Attachment.

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 ( $N=33,509$  at time of decision and  $N=31,287$  in month 60 after decision). Dashed lines show 90% confidence intervals.

Being employed before the trial seems to affect the results

We see that those who had been previously employed have employment decreases due to incarceration

There is evidence that those who were non-employed before the trial experience an increase in employment due to incarceration

In other results, the authors find that those previous non-employed were more likely to attend job training programs while in prison

# What Does it Mean to be Incarcerated in Norway?

- Their result that incarceration reduced recidivism and has mixed effects on employment may seem odd when thinking about incarceration in the US
- Prisons in Norway are different, where “life inside will resemble life outside as much as possible” and “offenders shall be placed in the lowest possible security regime”
- Low-level offenders to go “open” prisons, which are more like dorms
- All prisons offer education, mental health and training programs
- The most common programs are for high school and work-related training. Those not enrolled must work within prison

# What Does it Mean to be Incarcerated in Norway?

- Inmates have the right to daily physical exercise and access to a library and newspapers
- They have the same rights to health care as the regular population
- 18% of inmates participate in a drug-related program while in prison
- After release, there is also an emphasis on helping offenders reintegrate into society, with access to active labor market and other programs set up to help ex-convicts find a job and access social services like housing support

## ? Key question

- If this study were to have been done in the US, would the results be the same? Or would they be different since prisons in the US provide a much worse experience?



# Eren, Mocan (2021)



# Introduction to Eren, Mocan (2021)

- The study examines the effect of juvenile crime punishment on high school completion and adult recidivism in Louisiana
- Cases were randomly assigned to judges, allowing for estimation of incarceration's causal effects based on varying judge stringency
- Juvenile incarceration is linked to a higher likelihood of adult drug offense convictions but a lower likelihood of property crime convictions
- A negative impact of juvenile incarceration on high school completion was observed for earlier cohorts, with no significant effect on later cohorts

# Summary Statistics

Table 1: Summary Statistics

|                                         | Full Sample |       | Incarcerated |       | Non Incarcerated |       |
|-----------------------------------------|-------------|-------|--------------|-------|------------------|-------|
|                                         | Mean        | SD    | Mean         | SD    | Mean             | SD    |
|                                         | (1)         | (2)   | (3)          | (4)   | (5)              | (6)   |
| Panel A: Juvenile Characteristics       |             |       |              |       |                  |       |
| Incarcerated as a Juvenile              | 0.247       | 0.431 | 1.000        | 0.000 | 0.000            | 0.000 |
| Black                                   | 0.653       | 0.476 | 0.745        | 0.436 | 0.623            | 0.485 |
| White                                   | 0.328       | 0.469 | 0.234        | 0.423 | 0.360            | 0.480 |
| Female                                  | 0.252       | 0.435 | 0.122        | 0.328 | 0.296            | 0.456 |
| Age at Conviction                       | 15.09       | 1.35  | 15.44        | 1.18  | 14.98            | 1.38  |
| Juvenile Offense Type:                  |             |       |              |       |                  |       |
| Drug Related                            | 0.123       | 0.329 | 0.155        | 0.362 | 0.113            | 0.317 |
| Violent                                 | 0.081       | 0.273 | 0.153        | 0.360 | 0.058            | 0.234 |
| Property                                | 0.389       | 0.487 | 0.411        | 0.492 | 0.382            | 0.486 |
| Other                                   | 0.406       | 0.491 | 0.280        | 0.449 | 0.446            | 0.497 |
| Panel B: Adult Characteristics/Outcomes |             |       |              |       |                  |       |
| Adult Conviction                        | 0.387       | 0.487 | 0.547        | 0.498 | 0.335            | 0.472 |
| Adult Crime Type:                       |             |       |              |       |                  |       |
| Drug Related                            | 0.163       | 0.369 | 0.218        | 0.413 | 0.145            | 0.352 |
| Violent                                 | 0.068       | 0.253 | 0.116        | 0.320 | 0.053            | 0.224 |
| Property                                | 0.139       | 0.346 | 0.194        | 0.396 | 0.121            | 0.326 |
| Other                                   | 0.031       | 0.173 | 0.042        | 0.200 | 0.028            | 0.164 |
| Age of Adult Crime                      | 19.77       | 2.19  | 19.43        | 2.00  | 19.95            | 2.27  |
| Graduated High School                   | 0.238       | 0.426 | 0.167        | 0.373 | 0.261            | 0.439 |
| Sample Size                             | 7,371       |       | 1,822        |       | 5,549            |       |

NOTES: The statistics above reflect our research sample, which consists of one-time juvenile offenders over a period from 1996 to 2004 who were 25 years or younger by 2012 (birth cohorts between 1979 and 1987). The sample is further restricted to juveniles whose disposition decisions are made in courts where there were at least two regular judges in a given year (1996-2004).

The juvenile defendants in their data are mostly black (65.3%), male (74.8%), and are age 15 on average

Property and “other” crimes are most common

On average 38.7% of these juvenile defendants get a conviction in adulthood, with the average age of the adult crime being at almost age 20

Only 23.8% of the juvenile defendants will go on to eventually graduate high school

# First Stage: Testing the Stringency of Judges

Table 2: First Stage Results-The Effect of Judge Stringency in Incarceration on Juvenile Incarceration

|                                         | Juvenile Incarceration            |                     |                     |
|-----------------------------------------|-----------------------------------|---------------------|---------------------|
|                                         | Coefficients<br>(Standard Errors) |                     |                     |
|                                         | (1)                               | (2)                 | (3)                 |
| Judge Stringency in Incarceration       | 0.798***<br>(0.179)               | 0.755***<br>(0.177) | 0.814***<br>(0.149) |
| <i>F-Stat</i>                           | 19.79                             | 18.22               | 29.81               |
| Sample Size                             | 7,371                             | 7,371               | 7,371               |
| Controls:                               |                                   |                     |                     |
| Court-by-Disposition Year Fixed Effects | Yes                               | Yes                 | Yes                 |
| Juvenile                                | No                                | Yes                 | Yes                 |
| Juvenile Offense Fixed Effects          | No                                | No                  | Yes                 |

NOTES: Standard errors, which are clustered at the judge level, are reported in parentheses. There are 73 judges in total. Juvenile controls include indicators for juvenile's gender and race as well as age and its square. There are 136 detailed offense types in the effective sample. Judge stringency is the leave-one-out mean incarceration rate obtained using all case files (past and future over a period from 1996 to 2012) a judge has handled (for judges with at least 25 case files).

\* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

First stage = does judge stringency affect incarceration?

Answer: Yes, absolutely. Incarceration rates are strongly linked to judge stringency



# Testing the Random Assignment of Judges

Table 3: Randomization Tests

|                                       | Judge Stringency in Incarceration |
|---------------------------------------|-----------------------------------|
|                                       | Coefficients<br>(Standard Errors) |
|                                       | (1)                               |
| Black                                 | 0.0008<br>(0.0013)                |
| White                                 | -0.0013<br>(0.0013)               |
| Female                                | -0.0005<br>(0.0009)               |
| Age of Juvenile Offense Conviction    | 0.0003<br>(0.0003)                |
| Juvenile Offense Type:                |                                   |
| Drug Related                          | -0.0004<br>(0.0010)               |
| Violent                               | -0.0023<br>(0.0017)               |
| Property                              | -0.0007<br>(0.0009)               |
| Felony                                | -0.0025*<br>(0.0014)              |
| Joint Significance ( <i>p-value</i> ) | 0.27                              |
| Sample Size                           | 7,371                             |

NOTES: Each cell represents a separate regression and all regression estimations control for court-by-disposition year fixed effects. Standard errors, which are clustered at the judge level, are reported in parentheses. See also notes to Table 2 and the text for further details.

\* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

Is assignment to more lenient or more stringent judges as-good-as-random?

Answer: Yes, there is almost no association between defendant characteristics and if they were assigned to a more stringent judge

(Note: one estimate is significant at the 10% level, but even if there are no effects for all these variables, we would expect to find a false positive 10% of the time. So, likely this one significant estimate is a false positive)

# Results: The Effect of Incarceration on Recidivism

Table 6: IV and Reduced Form Results- The Effect of Juvenile Incarceration on Adult Criminal Convictions

|                                         | IV Results                        |                      |                      | Reduced Form                         |                      |
|-----------------------------------------|-----------------------------------|----------------------|----------------------|--------------------------------------|----------------------|
|                                         | Coefficients<br>(Standard Errors) |                      |                      |                                      |                      |
|                                         | (1)                               | (2)                  | (3)                  |                                      | (4)                  |
| Panel A: Any Crime                      |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | 0.049<br>(0.203)                  | 0.016<br>(0.185)     | 0.013<br>(0.160)     | Judge Stringency<br>in Incarceration | 0.010<br>(0.129)     |
| Panel B: Drug Related Crimes            |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | 0.302**<br>(0.138)                | 0.290**<br>(0.137)   | 0.276**<br>(0.119)   | Judge Stringency<br>in Incarceration | 0.225***<br>(0.077)  |
| Panel C: Violent Crimes                 |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | -0.017<br>(0.088)                 | -0.026<br>(0.086)    | -0.027<br>(0.076)    | Judge Stringency<br>in Incarceration | -0.022<br>(0.062)    |
| Panel D: Property Crimes                |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | -0.412***<br>(0.110)              | -0.441***<br>(0.109) | -0.413***<br>(0.092) | Judge Stringency<br>in Incarceration | -0.335***<br>(0.059) |
| Sample Size                             | 7,371                             | 7,371                | 7,371                |                                      | 7,371                |
| Controls:                               |                                   |                      |                      |                                      |                      |
| Court-by-Disposition Year Fixed Effects | Yes                               | Yes                  | Yes                  |                                      | Yes                  |
| Juvenile                                | No                                | Yes                  | Yes                  |                                      | Yes                  |
| Juvenile Offense Fixed Effects          | No                                | No                   | Yes                  |                                      | Yes                  |

NOTES: Standard errors, which are clustered at the judge level, are reported in parentheses. There are 73 judges in total. Juvenile controls include indicators for juvenile's gender and race as well as age and its square. There are 136 detailed offense types in the effective sample. Adult crime takes the value of one if juvenile is convicted as adult at age 25 or younger. See also notes to Table 2 and the text for further details.

\* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

Using judge fixed effects, where they compare on-average juvenile defendants assigned either more lenient or more picky judges, they find that

Incarceration increases drug-related crimes

Incarceration decreases property crimes

# Results: The Effect of Incarceration on Recidivism

Table 6: IV and Reduced Form Results- The Effect of Juvenile Incarceration on Adult Criminal Convictions

|                                         | IV Results                        |                      |                      | Reduced Form                         |                      |
|-----------------------------------------|-----------------------------------|----------------------|----------------------|--------------------------------------|----------------------|
|                                         | Coefficients<br>(Standard Errors) |                      |                      |                                      |                      |
|                                         | (1)                               | (2)                  | (3)                  |                                      | (4)                  |
| Panel A: Any Crime                      |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | 0.049<br>(0.203)                  | 0.016<br>(0.185)     | 0.013<br>(0.160)     | Judge Stringency<br>in Incarceration | 0.010<br>(0.129)     |
| Panel B: Drug Related Crimes            |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | 0.302**<br>(0.138)                | 0.290**<br>(0.137)   | 0.276**<br>(0.119)   | Judge Stringency<br>in Incarceration | 0.225***<br>(0.077)  |
| Panel C: Violent Crimes                 |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | -0.017<br>(0.088)                 | -0.026<br>(0.086)    | -0.027<br>(0.076)    | Judge Stringency<br>in Incarceration | -0.022<br>(0.062)    |
| Panel D: Property Crimes                |                                   |                      |                      |                                      |                      |
| Juvenile Incarceration                  | -0.412***<br>(0.110)              | -0.441***<br>(0.109) | -0.413***<br>(0.092) | Judge Stringency<br>in Incarceration | -0.335***<br>(0.059) |
| Sample Size                             | 7,371                             | 7,371                | 7,371                |                                      | 7,371                |
| Controls:                               |                                   |                      |                      |                                      |                      |
| Court-by-Disposition Year Fixed Effects | Yes                               | Yes                  | Yes                  |                                      | Yes                  |
| Juvenile                                | No                                | Yes                  | Yes                  |                                      | Yes                  |
| Juvenile Offense Fixed Effects          | No                                | No                   | Yes                  |                                      | Yes                  |

NOTES: Standard errors, which are clustered at the judge level, are reported in parentheses. There are 73 judges in total. Juvenile controls include indicators for juvenile's gender and race as well as age and its square. There are 136 detailed offense types in the effective sample. Adult crime takes the value of one if juvenile is convicted as adult at age 25 or younger. See also notes to Table 2 and the text for further details.

\* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

In other results, these effects are stronger the longer the person is in prison, especially for drug-related crimes

In other results, for those born before 1983, incarceration reduces the probability of high school graduation

# Summary of Bhuller et al. (2016) and Eren and Mocan (2019)

- Bhuller et al. (2016) – Adults, Norway – Incarceration reduces recidivism, mixed impacts on labor market outcomes
- Eren and Mocan (2019) – Juveniles, Louisiana – Incarceration reduces property crime recidivism but increases drug crime recidivism. Incarceration also reduces the probability of high school graduation for earlier cohorts (born before 1983)
- Are the different results due to the different contexts? (e.g., juveniles vs. adults, Louisiana prisons vs. Norway prisons)
- More research is likely needed. These are two recent new papers that get around the difficult “selection” issue by using administrative data and judge fixed effects

# Overview of Audit Studies



# What are Audit Studies?

- The most common way that economists and sociologists measure discrimination
- Send on-average identical “testers” (e.g., resumes, emails, actors) that vary by minority status (e.g., white vs. black) to study discrimination in hiring, market access, etc
- Since the testers only differ by race, any differences in responses to them can isolate discrimination
- Audit field experiments are commonly used to study housing discrimination

# Audit field experiments - Housing

- Phillips (2016) finds housing rental ads online (e.g., Craigslist)
- He then sends housing rental viewing requests from potential tenants to these landlords that posted the rental ads
- Viewing requests comes from individuals with white-sounding name (e.g., Emily Smith) of African-American names (e.g., Lakisha Washington)
- Phillips (2016) also randomly includes a mention of if the individual is going to pay with a Section 8 voucher
  - This is an anti-poverty program there the government pays part of your rent

# Audit field experiments - Housing

**Table 3**  
Effect of voucher message on the probability of response.

|                      | (1)<br>Positive      | (2)<br>Positive      | (3)<br>Positive      | (4)<br>Positive     | (5)<br>Any           | (6)<br>Showing       |
|----------------------|----------------------|----------------------|----------------------|---------------------|----------------------|----------------------|
| Section 8            | -0.27***<br>(0.021)  | -0.26***<br>(0.021)  | -0.26***<br>(0.021)  | -0.26***<br>(0.037) | -0.15***<br>(0.036)  | -0.18***<br>(0.030)  |
| Black Name           | -0.061***<br>(0.017) | -0.062***<br>(0.017) | -0.062***<br>(0.017) | -0.060**<br>(0.025) | -0.085***<br>(0.025) | -0.051**<br>(0.024)  |
| Positive Signal      |                      | 0.017<br>(0.023)     | 0.019<br>(0.023)     | 0.035<br>(0.035)    | 0.024<br>(0.035)     | 0.057*<br>(0.032)    |
| Negative Signal      |                      | -0.17***<br>(0.023)  | -0.17***<br>(0.023)  | -0.16***<br>(0.037) | -0.066*<br>(0.035)   | -0.088***<br>(0.032) |
| Apt. FE              | N                    | N                    | N                    | Y                   | Y                    | Y                    |
| Apt. Characteristics | N                    | N                    | Y                    | N                   | N                    | N                    |
| Mean of Dep. Var.    | 0.43                 | 0.43                 | 0.43                 | 0.43                | 0.59                 | 0.26                 |
| R-Squared            | 0.057                | 0.088                | 0.12                 | 0.73                | 0.74                 | 0.73                 |
| No. Obs.             | 2681                 | 2681                 | 2655                 | 2681                | 2681                 | 2681                 |

All estimates result from OLS estimation of a linear probability model. Standard errors clustered by apartment are in parentheses.

\* Indicate statistical significance at the 10% levels.

\*\* Indicate statistical significance at the 5% levels.

\*\*\* Indicate statistical significance at the 1% levels.

The average positive response rate (e.g., viewing offer) is 43%

Those who mention Section 8 have about a 26 percentage point lower positive response rate (so, about 17%)

Those with African-American names have about a 6 percentage point lower positive response rate (so, about 37%)



# Audit field experiments - Hiring

Audit field experiments are commonly used to study hiring discrimination

These are the “resume experiments” you have likely heard of

The Agan and Starr (2018) paper, the focus of today, is one of these resume experiments, but with an interesting difference-in-differences twist (explained later)



# Audit field experiments - Hiring

Researchers sent on-average identical applications (resumes) to job ads

Applicants vary in minority status (e.g., white vs. African-American name, male vs. female, older vs. younger, criminal record vs. no criminal record)

Discrimination is quantified by comparing “callback rates” – interview or similar positive response rates by employers



# Ban the Box



# What is Ban the Box?

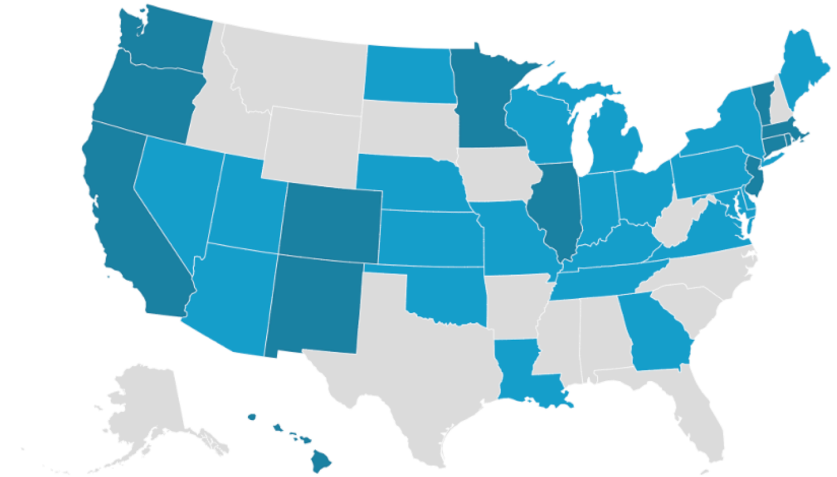
- Before I get into the methodology of the Agan and Starr (2018) resume experiment, I first want to explain what “Ban the Box” is
- From Wikipedia: **“Ban the Box is the name of an American campaign by advocates for ex-offenders, aimed at removing the check box that asks if applicants have a criminal record from hiring applications. Its purpose is to enable ex-offenders to display their qualifications in the hiring process before being asked about their criminal records. The premise of the campaign is that anything that makes it harder for ex-offenders to find a job makes it likelier that they will re-offend, which is bad for society.”**

# What is Ban the Box? 2019 Data

- Dark blue = box is banned for private employers and public employers
- Light blue = box is banned for public employers only
- Gray = no ban the box
- Note: this figure doesn't show city or county-level laws, just state laws
- There is no “Ban the Box” law at the federal level

## Statewide 'Ban the Box' Laws

States in blue prevent the government from screening job applicants for criminal history. Darker blue indicates the policy extends to private employers, while states in gray do not have either.



Source: National Employment Law Project • [Get the data](#)

# Agan and Starr (2018)



# Introduction to Agan and Starr (2018)

- "Ban the Box" (BTB) policies aim to help reduce unemployment among black men by restricting employer inquiries into applicants' criminal histories.
- There's a concern that BTB may lead to increased racial discrimination, with employers potentially making assumptions based on race.
- A study involving around 15,000 online job applications was conducted to assess the effects of BTB policies in New Jersey and New York City.
- Applications simulated young, male applicants with names signaling either black or white racial identity, along with varied felony conviction status.
- Evidence showed criminal records greatly reduce employment opportunities: employers without the box were 63% more likely to respond to applicants without a record.
- The racial gap in employer callbacks increased post-BTB, with white applicants initially receiving 7% more callbacks than black applicants, which widened to 43% after BTB.
- This suggests employers may rely on racial biases regarding criminality in the absence of explicit information about criminal records.

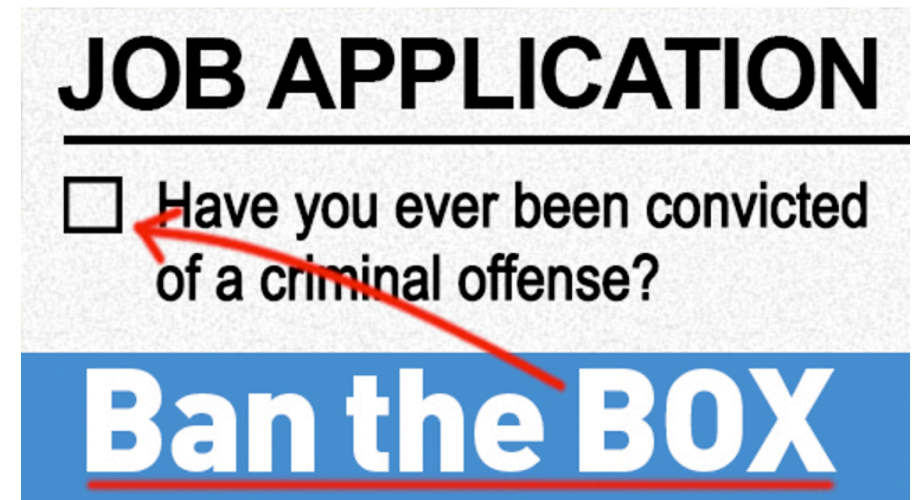
# President Obama on Ban the Box





# BTB Audit Study

- Agan and Starr (2018) study the impact of the enactment of a BTB law on hiring discrimination against those with criminal records and African-Americans
- They use an audit study – a resume field experiment – to quantify this discrimination
- Race or ethnicity is signaled through names (e.g., Greg Nelson, Darius Washington)
- Criminal record is signaled by checking the “box” or not, if it is available



**JOB APPLICATION**

☐ Have you ever been convicted of a criminal offense?

**Ban the BOX**

# BTB Audit Study

- This study is innovative because it uses an audit field experiment combined with a difference-in-differences to quantify the effect of BTB
- They first do a cross-sectional analysis, which just uses data before the BTB policy was passed
- They compare discrimination in jobs with and without a box, to see if having a box is associated with a different amount of racial discrimination
- They then use temporal variation in box, which occurs after NYC passes a BTB policy, and almost all NYC employers get rid of the box
- They compare New York City before and after its BTB to New Jersey during a similar time period, to see how racial discrimination in NYC changed during that time period, compared to the “control” of NJ

# Summary Statistics

TABLE I  
MEANS OF APPLICANT AND APPLICATION CHARACTERISTICS AND  
CALLBACK RATES BY PERIOD

|                                  | Pre-BTB | Post-BTB | Combined |
|----------------------------------|---------|----------|----------|
| Characteristics                  |         |          |          |
| White                            | 0.502   | 0.497    | 0.500    |
| Conviction                       | 0.497   | 0.513    | 0.505    |
| GED                              | 0.498   | 0.502    | 0.500    |
| Employment gap                   | 0.492   | 0.504    | 0.498    |
| Application has box              | 0.366   | 0.036    | 0.199    |
| Results                          |         |          |          |
| Callback rate                    | 0.109   | 0.125    | 0.117    |
| Interview req.                   | 0.060   | 0.067    | 0.063    |
| Callback rate by characteristics |         |          |          |
| Black                            | 0.099   | 0.111    | 0.105    |
| White                            | 0.120   | 0.139    | 0.129    |
| GED                              | 0.106   | 0.127    | 0.117    |
| HSD                              | 0.113   | 0.122    | 0.118    |
| Emp. gap                         | 0.110   | 0.126    | 0.118    |
| No emp. gap                      | 0.109   | 0.124    | 0.116    |
| N                                | 7,245   | 7,392    | 14,637   |

*Notes.* "Callback" means application received a personalized positive response from the employer (either voicemail or e-mail). Interview request means the positive response message specifically mentioned an interview. "Application has box" means that the application asked about criminal records. "Employment gap" or "Emp. gap" means an 11–13-month employment gap in work history, while "No emp. gap" means a 0–2-month gap.

- This table presents what their data looks like
- About half of the job applicants are white (vs. black), have a conviction (vs. no conviction), have a GED (vs. no GED), and have an 11-13 month employment gap in work history (vs. 0-2 month gap)
- Callback rates on average are 10.9%, and interview offer rates are 6.0%
  - The difference between them – 4.9%, is other positive employer response that is not an explicit interview offer

# Summary Statistics

TABLE I  
MEANS OF APPLICANT AND APPLICATION CHARACTERISTICS AND  
CALLBACK RATES BY PERIOD

|                                  | Pre-BTB | Post-BTB | Combined |
|----------------------------------|---------|----------|----------|
| Characteristics                  |         |          |          |
| White                            | 0.502   | 0.497    | 0.500    |
| Conviction                       | 0.497   | 0.513    | 0.505    |
| GED                              | 0.498   | 0.502    | 0.500    |
| Employment gap                   | 0.492   | 0.504    | 0.498    |
| Application has box              | 0.366   | 0.036    | 0.199    |
| Results                          |         |          |          |
| Callback rate                    | 0.109   | 0.125    | 0.117    |
| Interview req.                   | 0.060   | 0.067    | 0.063    |
| Callback rate by characteristics |         |          |          |
| Black                            | 0.099   | 0.111    | 0.105    |
| White                            | 0.120   | 0.139    | 0.129    |
| GED                              | 0.106   | 0.127    | 0.117    |
| HSD                              | 0.113   | 0.122    | 0.118    |
| Emp. gap                         | 0.110   | 0.126    | 0.118    |
| No emp. gap                      | 0.109   | 0.124    | 0.116    |
| N                                | 7,245   | 7,392    | 14,637   |

*Notes.* "Callback" means application received a personalized positive response from the employer (either voicemail or e-mail). Interview request means the positive response message specifically mentioned an interview. "Application has box" means that the application asked about criminal records. "Employment gap" or "Emp. gap" means an 11–13-month employment gap in work history, while "No emp. gap" means a 0–2-month gap.

- This table presents what their data looks like
- Callback rates are lower for Black people, slightly lower for GED compared to high school diploma, but are similar for those with and without employment gaps
- Callback rates seem to be increasing over time, hence the general increase comparing pre to post-BTB
- But the key question is if discrimination changes differentially after BTB (NYC) compared to the same time period without a BTB change (NJ)

# Callback Rates are Lower for Those with Criminal Records

TABLE II  
CALLBACK RATES BY CRIME STATUS FOR STORES WITH THE BOX IN THE PRE-BTB PERIOD

|                | No crime | Crime | Property | Drug  | Combined |
|----------------|----------|-------|----------|-------|----------|
| Callback rate  | 0.136    | 0.085 | 0.084    | 0.085 | 0.110    |
| Callback black | 0.131    | 0.086 | 0.091    | 0.081 | 0.109    |
| Callback white | 0.140    | 0.083 | 0.077    | 0.089 | 0.111    |
| <i>N</i>       | 1,319    | 1,336 | 703      | 633   | 2,655    |

*Notes.* Sample restricted to pre-BTB period applications where the application asked about criminal records. Callback implies application received a personalized positive response from the employer.

# Callback Rates are Lower for Those with Criminal Records

TABLE III  
EFFECTS OF APPLICANT CHARACTERISTICS ON CALLBACK RATES

|                  | (1)                 | (2)                  | (3)                  |
|------------------|---------------------|----------------------|----------------------|
| White            | 0.024***<br>(0.006) | −0.001<br>(0.009)    | −0.001<br>(0.009)    |
| Conviction       | −0.014**<br>(0.005) | −0.052***<br>(0.012) |                      |
| GED              | −0.004<br>(0.005)   | 0.010<br>(0.014)     | 0.010<br>(0.013)     |
| Employment gap   | 0.002<br>(0.005)    | 0.011<br>(0.010)     | 0.011<br>(0.010)     |
| Pre-BTB period   | −0.015<br>(0.010)   |                      |                      |
| Drug conviction  |                     |                      | −0.050***<br>(0.013) |
| Prop. conviction |                     |                      | −0.054***<br>(0.014) |
| <i>N</i>         | 14,637              | 2,918                | 2,918                |
| Sample           | All                 | Box                  | Box                  |
| Chain FE         | Yes                 | Yes                  | Yes                  |
| Center FE        | Yes                 | Yes                  | Yes                  |

*Notes.* Dependent variable is whether the application received a callback. Standard errors clustered on chain in parentheses. Chain and geographic center fixed effects are included in all regressions. White is an indicator for race (versus black), Conviction is an indicator for whether the applicant has a felony conviction, GED is an indicator for having a GED (versus a regular high-school diploma), and Employment gap is an indicator for whether the applicant has an 11–13-month gap in work history between the previous two jobs (versus a 0–2-month gap). “Drug conviction” and “Prop. conviction” break the Conviction variable into a categorical variable based on crime type (drug versus property crime); no conviction is the base category. The box sample is employers with the box on their application. \*10%, \*\*5%, and \*\*\*1% significance level.

# Difference-in-differences in Regression Form

$$Callback_{ij} = \alpha + \beta_1 Box_j + \beta_2 White_i + \beta_3 Box_j \times White_i + \Gamma \mathbf{X}_i$$

- They first use cross-sectional data (i.e. data from just one time period, in this case, pre BTB)
- They quantify how racial discrimination differs between firms that use or do not use the box
- The coefficient tells us the callback difference in general, regardless of race when there is a box
- The coefficient tells us the callback difference by race, regardless of if there is a box
- The coefficient gives the difference in differences estimate, which tells us if having a box differentially affects White vs Black people
- It captures the interaction effect: does having a box increase or decrease racial discrimination?

# Raw Data: Before BTB

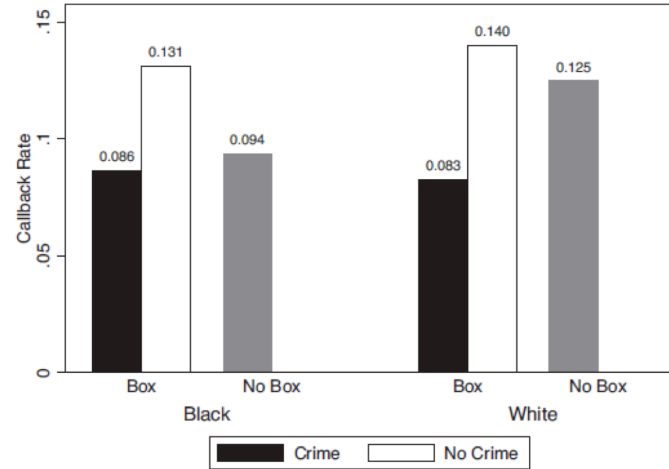


FIGURE I

Callback Rates by Race, Crime, and Box: Preperiod Applications Only

This figure compares callback rates within the preperiod before Ban the Box went into effect, comparing applications with the criminal record question box and those without the box.

- Callback rates are similar between White and Black people with a criminal record (only disclosed when there is a box)
- White people get a 0.9 percentage point higher callback rate compared to Black people when neither has a criminal record
- **Callback rates for White people are 3.1 percentage points higher, compared to Black people, when there is no box**



# Adding Temporal Variation

- Agan and Starr (2018) then build up to their ideal DiD specification by using temporal variation in BTB
- New York City passes a BTB law, which creates variation in BTB that can be used in a DiD to try to quantify the effect of BTB on discrimination
- This DiD is preferred to their earlier cross-sectional approach (similar to the naïve approach of only using one time period, from earlier)

# Adding Temporal Variation

- The DiD is preferred because there could be non-random reasons why some employers use a box and others don't. Given then, it's hard to isolate how the box affects discrimination compared to how firms that use the box differ and how this could affect racial differences
- For example, what if firms that use the box are more racist anyways? Then we might confuse the effect of the box for selection into who uses the box
- In this hypothetical example, banning the box doesn't reduce racial discrimination

# Results

TABLE IV  
EFFECTS OF THE BOX ON RACIAL DISCRIMINATION: DIFFERENCE-IN-DIFFERENCES

|                           | (1)           | (2)                         | (3)                     | (4)                     | (5)                     |
|---------------------------|---------------|-----------------------------|-------------------------|-------------------------|-------------------------|
| Box × white               | −0.030**      | −0.036**                    | −0.033**                | −0.027**                | 0.002                   |
| (White × pre, column (5)) | (0.015)       | (0.014)                     | (0.014)                 | (0.013)                 | (0.014)                 |
| White                     | 0.032***      | 0.044***                    | 0.040***                | 0.123                   | 0.022**                 |
|                           | (0.012)       | (0.013)                     | (0.012)                 | (0.132)                 | (0.009)                 |
| Box                       | 0.015         | 0.003                       | −0.002                  | −0.345**                | −0.016                  |
| (Pre, column (5))         | (0.024)       | (0.015)                     | (0.013)                 | (0.139)                 | (0.017)                 |
| N                         | 7,245         | 3,712                       | 4,794                   | 4,794                   | 7,476                   |
| Controls                  | Yes           | Yes                         | Yes                     | Yes                     | Yes                     |
| Center FE                 | Yes           | No                          | Yes                     | Yes                     | No                      |
| Chain FE                  | No            | No                          | No                      | Yes                     | No                      |
| Post × chain FE           | No            | No                          | No                      | Yes                     | No                      |
| White × chain FE          | No            | No                          | No                      | Yes                     | No                      |
| Box variation             | Cross-section | Temporal                    | Temporal                | Temporal                | None                    |
| Sample                    | Pre-BTB       | Box<br>remover<br>-balanced | Box<br>remover<br>-full | Box<br>remover<br>-full | Other empl.<br>balanced |

Notes. Standard errors clustered on chain in parentheses. Dependent variable is whether the application received a callback. Box removers are stores that had the box in the pre-BTB period and removed it after BTB. “Box removers-balanced” consists of box remover stores to which we sent exactly four applications, one white/black pair in each period. Fixed effects can include geographic center, chain, post × chain, and white × chain, and are included as indicated; note that because of the inclusion of interacted fixed effects in column (4), the white and box coefficients are not meaningful. Controls are whether the applicant had a GED (versus regular high-school diploma) and whether he had an employment gap. Box variation indicates the source of variation in the box variable: “Cross-section” means the variation comes from a comparison of box and nonbox stores in the preperiod; “Temporal” means the variation is pre- and post-BTB, triggered by the implementation of the BTB policy. In the last column, which is shown as a comparison point, there is no box variation; the pattern over the same time period is shown for companies that did not change their job applications. \*10%, \*\*5%, and \*\*\*1% significance level.

- Turns out their results are similar comparing the naïve cross-sectional approach (column (1)) to the more advanced DiD approach (columns (2) to (4))
- Looking at column (2), the coefficient on white shows a 4.4 percentage point higher callback rate for White people when there is no box
- The coefficient on box is not statistically significant – callback rates for White people are similar with and without the box

# Results

TABLE IV  
EFFECTS OF THE BOX ON RACIAL DISCRIMINATION: DIFFERENCE-IN-DIFFERENCES

|                           | (1)           | (2)                         | (3)                     | (4)                     | (5)                     |
|---------------------------|---------------|-----------------------------|-------------------------|-------------------------|-------------------------|
| Box × white               | −0.030**      | −0.036**                    | −0.033**                | −0.027**                | 0.002                   |
| (White × pre, column (5)) | (0.015)       | (0.014)                     | (0.014)                 | (0.013)                 | (0.014)                 |
| White                     | 0.032***      | 0.044***                    | 0.040***                | 0.123                   | 0.022**                 |
|                           | (0.012)       | (0.013)                     | (0.012)                 | (0.132)                 | (0.009)                 |
| Box                       | 0.015         | 0.003                       | −0.002                  | −0.345**                | −0.016                  |
| (Pre, column (5))         | (0.024)       | (0.015)                     | (0.013)                 | (0.139)                 | (0.017)                 |
| N                         | 7,245         | 3,712                       | 4,794                   | 4,794                   | 7,476                   |
| Controls                  | Yes           | Yes                         | Yes                     | Yes                     | Yes                     |
| Center FE                 | Yes           | No                          | Yes                     | Yes                     | No                      |
| Chain FE                  | No            | No                          | No                      | Yes                     | No                      |
| Post × chain FE           | No            | No                          | No                      | Yes                     | No                      |
| White × chain FE          | No            | No                          | No                      | Yes                     | No                      |
| Box variation             | Cross-section | Temporal                    | Temporal                | Temporal                | None                    |
| Sample                    | Pre-BTB       | Box<br>remover<br>-balanced | Box<br>remover<br>-full | Box<br>remover<br>-full | Other empl.<br>balanced |

Notes. Standard errors clustered on chain in parentheses. Dependent variable is whether the application received a callback. Box removers are stores that had the box in the pre-BTB period and removed it after BTB. “Box removers-balanced” consists of box remover stores to which we sent exactly four applications, one white/black pair in each period. Fixed effects can include geographic center, chain, post × chain, and white × chain, and are included as indicated; note that because of the inclusion of interacted fixed effects in column (4), the white and box coefficients are not meaningful. Controls are whether the applicant had a GED (versus regular high-school diploma) and whether he had an employment gap. Box variation indicates the source of variation in the box variable: “Cross-section” means the variation comes from a comparison of box and nonbox stores in the preperiod; “Temporal” means the variation is pre- and post-BTB, triggered by the implementation of the BTB policy. In the last column, which is shown as a comparison point, there is no box variation; the pattern over the same time period is shown for companies that did not change their job applications. \*10%, \*\*5%, and \*\*\*1% significance level.

- *Box* × *white*, the coefficient of interest is -0.036, suggesting that callback rates for White people are 3.6 percentage points lower, compared to Black people, when there is a box, compared to when there is no box
- Putting this together, the white “benefit” is 4.4 percentage points when there is no box, but only about 0.8 p.p.s when there is a box
- **Removing the box is associated with an increase in racial discrimination**

# Results

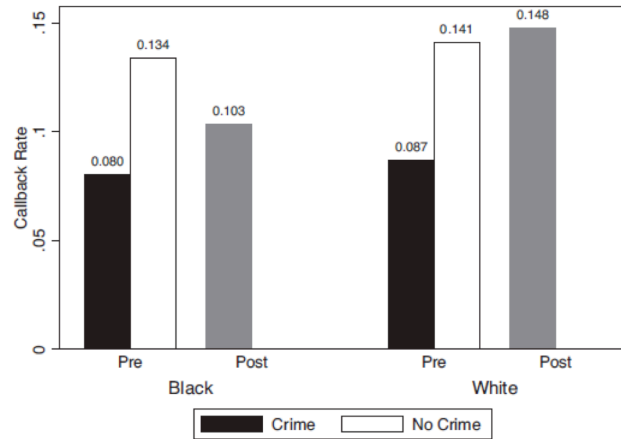


FIGURE II

Callback Rates by Race, Criminal Record, and Period: Balanced Box Removers Only

This figure compares callback rates before and after Ban the Box went into effect, among companies that had the criminal record question box before BTB and removed it afterward, in the balanced sample only (i.e., stores to which we sent complete application pairs in both the pre-BTB and post-BTB periods).

- The White “benefit” is about 7 percentage points when a box is present, regardless of criminal record
- **Callback rates for White people are 4.5 percentage points higher, compared to Black people, when there is no box**
- This is similar to the cross-sectional results, but a bigger magnitude of an effect

# Conclusion

- The evidence from this paper suggests that employers statistically discriminate against black applicants, by being more likely to assume they have criminal convictions when they do not have information on criminal conviction status
- When the box is banned, employers who used to rely on the box make the assumption that black applicants are more likely to have a criminal record
- Therefore, we see racial discrimination increase after BTB takes effects
- This suggests unintended consequences, as BTB was ideally supposed to reduce employment disparities for Black people, who are more likely to have criminal records
- There is the benefit that without a box, there is less discrimination against those with records, and those with records are disproportionately black, so that is still a benefit of BTB
- But the cost is an increase, in general, in discrimination against Black people