If You Give a Judge a Risk Score

Evidence from Kentucky Bail Decisions

Alex Albright, Harvard University

Eighth ECINEQ Meeting 2019

July 3, 2019

Predictive Scores and High-Stakes Decisions

Should we (loan officers) give you a loan?

$\rightarrow$ FICO scores

Should we (managers) hire you?

$\rightarrow$ job testing

How should we (pretrial judges) set your bond?

$\rightarrow$ risk assessment scores

Criminal Justice and Scores

Scores used in bail, pretrial, or sentencing hearings in 49 of the 50 US states (Traughber 2018)

- Think of score as a "check-list" instrument *("Add 1 point if person has a prior felony conviction")*
  - Variation over states and counties in what goes into them
- SB10 in CA to get rid of money bail and move to risk score system

Ongoing debate about implications of risk scores for racial justice

- ACLU opposed SB10 due to concerns about implementation's consequences on communities of color
- Eric Holder: score usage "may exacerbate unwarranted and unjust disparities"
- Michelle Alexander on "e-carceration": "The *Newest* Jim Crow"

Automation vs. Discretion

1. With automation, mechanically, people with same risk scores get the same treatment

2. In practice, human discretion in how to use the scores

Kentucky Judge (51st Judicial District):

> *Judges are people. When I have to defer to this mathematical model that I don't really understand all that well... it's hard. I'm going to do the best I can with what they've given me.*

Research Questions

How does requiring judges to use risk scores impact racial disparities?

What about for defendants with the same scores?

$\rightarrow$ Do judges follow risk score recommendations similarly across racial groups?

This Project

Policy change (June 2011) in Kentucky pretrial environment:
House Bill 463

1. Required judges to consider risk scores *(that were already in existence and optional)* in initial bond decisions

2. Made default for low/moderate defendants non-financial bond *(judges can deviate but must give reason)*

Qs: How does requiring judges to use risk scores impact racial disparities? (Building on Stevenson, 2017) What about for defendants with the same scores? Do judges follow risk score policy recommendations similarly across racial groups?

Roadmap

Literature

Discussion of risk score generation

- Angwin and Kirchner (2016), Kleinberg et al. (2017), Corbett-Davies et al. (2017), Yang and Dobbie (2019), Kleinberg et al. (2018)

Evidence that human decisions makers deviate from score recommendations

- Garrett and Monahan (2017), DeMichele et al. (2018), Hoffman, Kahn, and Li (2017), Main (2016), Skeem, Scurich, and Monahan (2019), Green and Chen (2019)

Risk score policy investigations

- Stevenson (2017), Sloan et al. (2018), Stevenson and Doleac (2019)

Kentucky Pretrial System (2009-2013)

$\rightarrow$ Imagine you're arrested and booked into jail by the police

$\rightarrow$ A judge makes an initial bond decision within 24 hours

$\rightarrow$ There are 3 steps to this

1. Pretrial Officer Collects Information

Pretrial officer (Pretrial Services employee) collects data on your arrest and charges

- police officer has full authority to charge; no prosecutorial review before bail decision
- looks up/physically collects data on defendant criminal history/offense
- interviews defendant

With all this data, they also calculate a **Kentucky Pretrial Risk Assessment level**

Kentucky Pretrial Risk Assessment (KPRA)

A checklist-style instrument:

- No verified local address: $+2$
- No verified means of support: $+1$
- A/B/C felony: $+1$
- New charge with pending case: $+7$
- FTA warrant or FTA misdemeanor or felony? $+2$
- Prior traffic FTA? $+1$
- Prior misdemeanor convictions? $+2$
- Prior felony convictions? $+1$
- Prior violent crime convictions? $+1$
- Drug/alcohol abuse? $+2$
- Conviction for felony escape? $+3$
- On probation/parole from felony conviction? $+1$

Low $= 0$-5; Moderate $= 6$-13; High $= 14$-max

2. Pretrial Officer-Judge Interaction

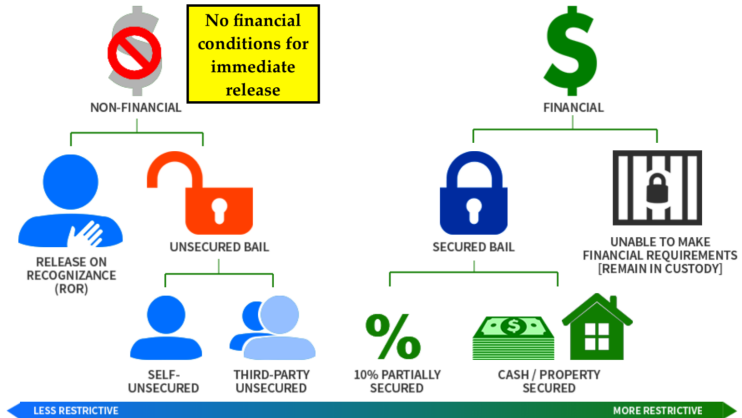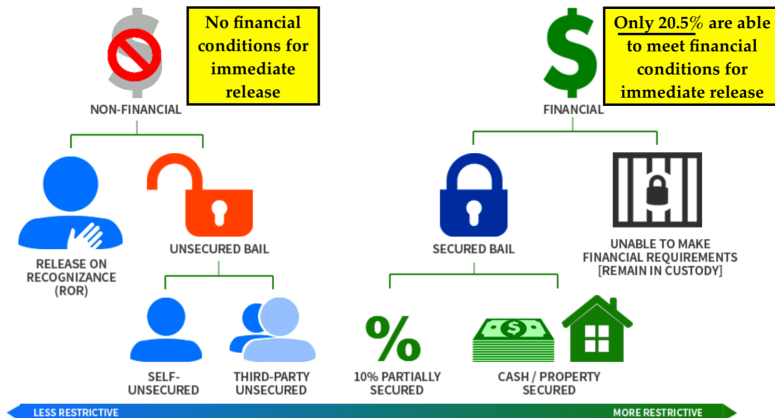# 3. Judge makes initial bond decision

3. Judge makes initial bond decision

3. Judge makes initial bond decision

Judicial Decision

NON-FINANCIAL

No financial conditions for immediate release

FINANCIAL

Only 20.5% are able to meet financial conditions for immediate release

RELEASE ON RECOGNIZANCE (ROR)

UNSECURED BAIL

SECURED BAIL

UNABLE TO MAKE FINANCIAL REQUIREMENTS [REMAIN IN CUSTODY]

SELF-UNSECURED

THIRD-PARTY UNSECURED

10% PARTIALLY SECURED

CASH / PROPERTY SECURED

LESS RESTRICTIVE   MORE RESTRICTIVE

3. Judge makes initial bond decision

Judicial Decision — Main outcome of interest: whether bond is non-financial or financial

Enter HB463

HB463 Background:

- Between "2000 and 2010, Kentucky's incarcerated population – both jail and prison – grew by 45%, more than three times the U.S. average." (Stevenson, 2017)
- HB463 passed in response to budget concerns (effective date 6/8/11)
  - Goal was to decrease pretrial detention

After HB463, again,

- Within 24 hours, pretrial officer collects data and makes a presentation to the judge
- Judges has a few minutes to make initial bond decision

But now. . .

Judge Interaction After HB463

KENTUCKY COURT OF JUSTICE

KENTUCKY PRETRIAL SERVICES VIRTUAL TOUR

Within 24 hours of booking:

Presentation to the Judge: Judge Call

- Typically handled through a **phone call**.
- Presents the defendant's **charge, risk-assessment score** and **recommendations on release**.
- Judges are **required to release low- and moderate-risk** defendants before trial on their own recognizance or unsecured bond.
- Judge must **document the reasons for denying release** in a written order. Reasons include that the defendant is found to be a **flight risk and/or a danger to the community**.

Presumptive default (for low/moderate) set by HB463
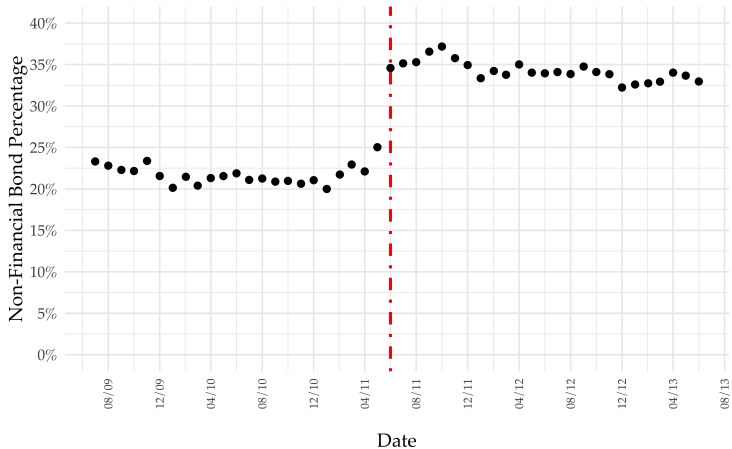Judges can deviate but must give a reason

Data

Data on 383,080 initial bond decisions for male defendants between 7/1/09-6/30/13

- Data via Kentucky Administrative Office of the Courts
- KPRA time period (changed to different scoring system 7/1/13)
- Spans 192,758 distinct defendants, 563 distinct judges
- 79.1% white, 20.6% black
- limiting to misdemeanors and felonies; 79.5% top charges are misdemeanors
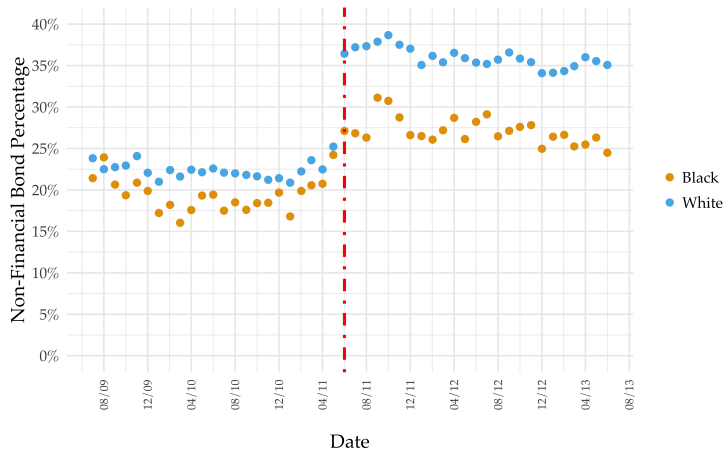- 68.3% financial, 27.8% non-financial bond, 3.9% no bond

Side note: Since being accepted to this conference, my understanding is that this sort of analysis of judge data has become illegal in France. (Article 33 of the Justice Reform Act)

Non-Financial Decisions Before/After HB463

Data from Kentucky AOC 7/1/09-6/30/13
Binned by month-year; red line marks the effective month of HB463
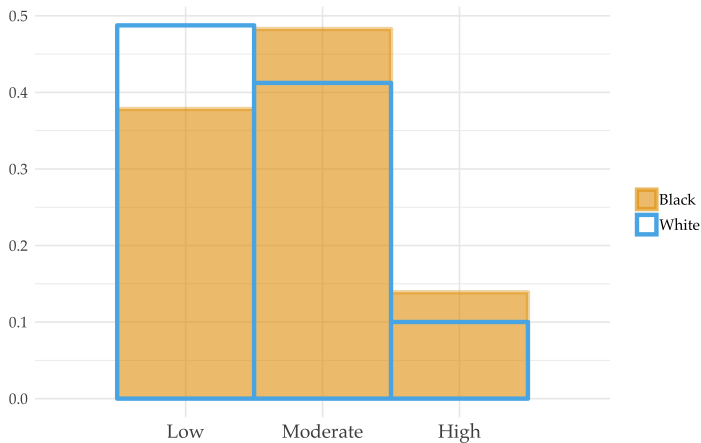
Does this look the same across rac...? No.

Data from Kentucky AOC 7/1/09-6/30/13
Binned by month-year; red line marks the effective month of HB463

Begs the questions. . .

Is the jump in racial disparities in non-financial bond a consequence of different risk levels?

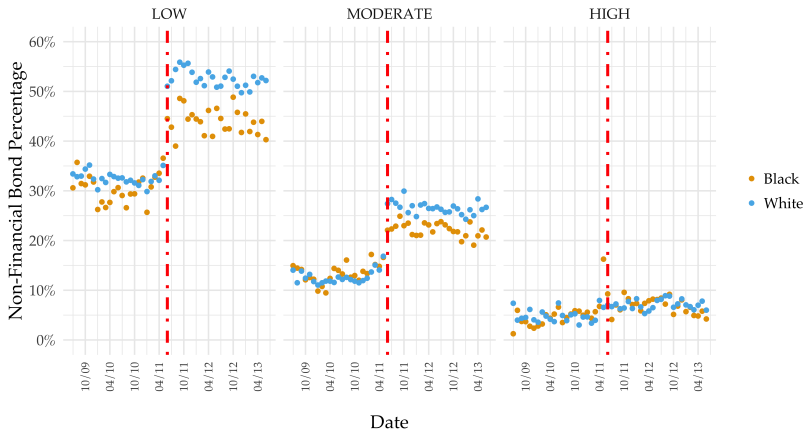Or is deviation from the presumptive default more likely for black defendants?

Risk level distributions look different by race

KPRA Level

Data from Kentucky 7/1/09-6/30/13

But risk levels don't explain the gaps

Data from Kentucky AOC 7/1/09-6/30/13
Binned by month-year; red line marks the effective month of HB463

Take-Away

The presumptive default (of non-financial bond for low and moderate risk defendants) is *more likely to be overridden for black defendants than white defendants*

What drives these differences in deviations across racial groups?

1. different underlying charges or defendant characteristics (different judge information sets)

2. different policy responses across judges

3. different treatment of similar defendants within judge and time

Why are there disparities in deviations?

Estimate three specifications *(raw, with judge info set, with judge info set and time-varying judge FEs)* for each risk level:

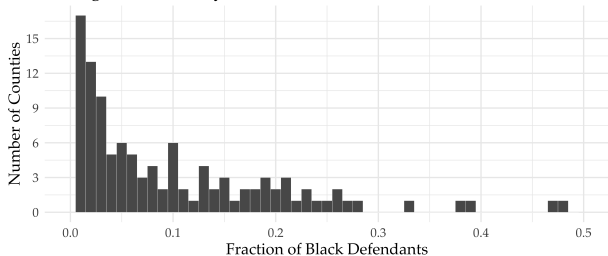$$b_{it} = \alpha + \phi_1 HB463_t + \phi_2 Black_i + \phi_3(Black_i \times HB463_t) + \epsilon_{it}$$

$$b_{ijct} = \alpha + \phi_1 HB463_t + \phi_2 Black_i + \phi_3(Black_i \times HB463_t) +$$
$$\beta_1 \kappa_c + \beta_2 \delta_i + \omega_t + x_t + \epsilon_{ijct}$$

$$b_{ijct} = \alpha + \phi_1 HB463_t + \phi_2 Black_i + \phi_3(Black_i \times HB463_t) +$$
$$\beta_1 \kappa_c + \beta_2 \delta_i + \omega_{jt} + \epsilon_{ijct}$$

With: b: dummy variable for non-financial bond; κ_c: vector of charge variables (severity, characteristics of offense); δ_i: vector of defendant variables (characteristics, criminal history variables); ω_j: judge FEs; x_t: time FEs; ω_{jt}: time-varying (month-year) judge FEs; ϕ_3: **coefficient of interest**
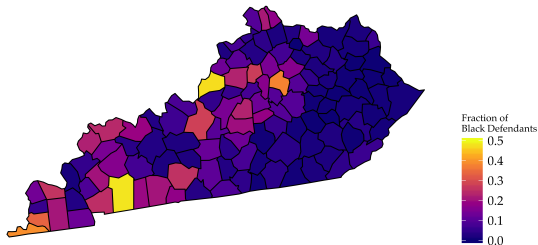
Why might judge responses be important? *(spoiler)*

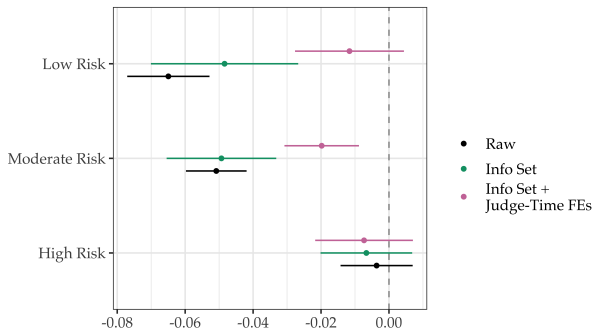Histogram of County-Level Fractions of Black Defendants

Binwidth is 0.01.

Choropleth of County-Level Fractions of Black Defendants

Data from Kentucky AOC 7/1/09-6/30/13

Results

Black x Post Coefficient Estimate
Whiskers depict 95% CIs
Data from Kentucky 7/1/09-6/30/13

- ▶ judge info sets don't meaningfully explain the changes in gaps
- ▶ low risk disparity changes mostly explained by judge-time FEs
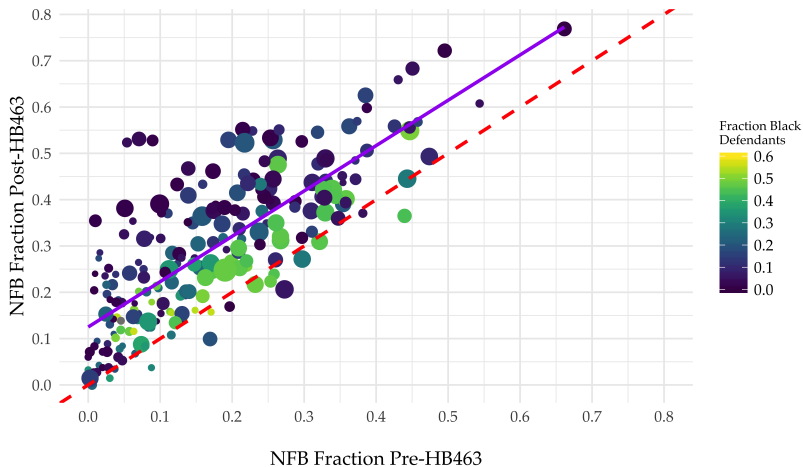- ▶ moderate risk disparity changes remain after inclusion of judge-time FEs

Take-Aways

(i) judges varied in their policy responsiveness; judges in whiter counties responded more to the new default than judges in blacker counties

(ii) suggestive evidence that interaction with the same predictive score may lead to different predictions by race (within judge)

Judicial Responsiveness Correlates with Population

Each point is one of 233 judges in Kentucky.
These are the judges who had at least 100 decisions in both time periods.
Area of points is scaled to the total number of decisions.
The red dashed line is a 45 degree line. The purple line is the regression line of post on pre fraction.
Data from Kentucky AOC 7/1/09-6/30/13

Future work: *why* this relationship?

Two hypotheses:

(1) judicial experience

larger counties in KY are higher % black, might be more prestigious to work there, more experienced judges get those roles, experience means you respond less to policy suggestions

(2) pretrial misconduct rates

judges who have made decisions with higher pretrial misconduct rates are less likely to respond strongly (more saturated), if misconduct rates are higher for judges with higher % black defendants, this could be a response to that

Moderate Risk Result

Result: within judge-time, racial disparity for moderate risk defendants after HB463 (but not before)

Possible explanation: judge interaction with the same predictive score level can lead to different predictions by race

- empirical evidence related to Kleinberg and Mullainathan (2019) theoretical result
  - simplified prediction functions (e.g., risk assessments) create incentives for decision-makers to consider group membership information
- accords with Green and Chen (2019) and Skeem, Scurich, and Monahan (2019)
- relevant to hiring, loan decisions, and other important high-stakes decisions

Conclusion

Results highlight the potential for:

(i) hetereogeneous judicial policy responses (correlated with geography) to generate exacerbated racial disparities in aggregate

(ii) risk score policies to generate disparate impacts even conditional on the scores themselves

Thank you!

Comments? Feedback?

Ping me: apalbright@g.harvard.edu

Full Results Table

		Low			Moderate			High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black	−0.022***	0.004	−0.011**	0.007**	0.022***	0.006*	0.002	0.005	0.006
	(0.004)	(0.006)	(0.005)	(0.003)	(0.005)	(0.003)	(0.004)	(0.005)	(0.005)
Post	0.201***	0.150***	0.145***	0.138***	0.079***	0.081***	0.021***	0.039*	0.042
	(0.003)	(0.020)	(0.021)	(0.002)	(0.020)	(0.025)	(0.003)	(0.020)	(0.027)
Black x Post	−0.065***	−0.048***	−0.012	−0.051***	−0.049***	−0.020***	−0.004	−0.007	−0.007
	(0.006)	(0.011)	(0.008)	(0.005)	(0.008)	(0.006)	(0.005)	(0.007)	(0.007)
Covariates?	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Judge-Time FEs?	No	No	Yes	No	No	Yes	No	No	Yes
Pre-White Mean	0.326	0.326	0.326	0.128	0.128	0.128	0.050	0.050	0.050
Cluster SE?	NA	Judge	Judge	NA	Judge	Judge	NA	Judge	Judge
N	178,238	178,238	178,238	163,479	163,479	163,479	41,363	41,363	41,363
R^2	0.039	0.220	0.294	0.027	0.144	0.231	0.002	0.068	0.299
Adjusted R^2	0.039	0.217	0.242	0.027	0.141	0.166	0.002	0.058	0.091

Dependent Variable = Non-Financial Bond

*** p<0.01; ** p<0.05; * p<0.1.