# The hidden effects of algorithmic recommendations

Alex Albright
Opportunity and Inclusive Growth Institute
Federal Reserve Bank of Minneapolis

# Algorithms in decision-making

- **Hiring:** resume scores
- **Consumer finance:** credit scores
- **Housing**: housing readiness
- **Health**: risk scores for mental health
- **Justice**: risk scores for pretrial misconduct

# Algorithms in decision-making

- **Hiring:** resume scores
- **Consumer finance:** credit scores
- **Housing**: housing readiness
- **Health**: risk scores for mental health
- **Justice**: risk scores for pretrial misconduct

But human decision-makers often make the final decision

*(e.g., hiring managers, therapists, judges)*

# How are algorithms given to decision-makers?

Consider a hiring manager reviewing a job applicant...

*Resume score: high*

Algorithm output
*(prediction from algorithm)*

# How are algorithms given to decision-makers?

Consider a hiring manager reviewing a job applicant...

| | | |
|---|---|---|
| *Resume score: high* | + | *Interview this person* |
| Algorithm output<br>*(prediction from algorithm)* | | Algorithmic recommendation<br>*(mapping from prediction to action)* |

# How are algorithms given to decision-makers?

Consider a hiring manager reviewing a job applicant…

*Resume score: high* + *Interview this person*

Algorithm output
*(prediction from algorithm)*

Algorithmic recommendation
*(mapping from prediction to action)*

**Studying “the effect of algorithms” on decisions often confounds these two components**

# How are algorithms given to decision-makers?

Consider a hiring manager reviewing a job applicant…

| *Resume score: high* | + | *Interview this person* |
|---|---|---|
| Algorithm output<br>*(prediction from algorithm)* | | Algorithmic recommendation<br>*(mapping from prediction to action)* |

**Studying “the effect of algorithms” on decisions often confounds these two components**

*Empirical challenge: usually introduced at the same time*

# The hidden effects of algorithmic recommendations

- Predictions generated by algorithms are distinct from algorithmic recommendations, which are often overlooked
    - **This paper: isolate the effects of recommendations**

# The hidden effects of algorithmic recommendations

- Predictions generated by algorithms are distinct from algorithmic recommendations, which are often overlooked
    - **This paper: isolate the effects of recommendations**

**How?** => Leverage a setting *(judges making bail decisions in CJS)* where

1. algorithmic predictions *(risk scores)* available to decision-makers stay the same
2. BUT algorithmic recommendations vary

# Preview of Results

1. Basic fact: **Recommendations matter**
    - Algorithmic recommendations impact decisions
      *(a lenient recommendation increases lenient bail for marginal cases by 50+%)*
    - Recommendations have independent effects from algorithms themselves

# Preview of Results

1. Basic fact: **Recommendations matter**
    - Algorithmic recommendations impact decisions
      *(a lenient recommendation increases lenient bail for marginal cases by 50+%)*
    - Recommendations have independent effects from algorithms themselves
2. Testing mechanisms: **Recommendations can change private costs of errors**
    - Making mistakes is less costly when decision consistent with recommendation
      *(lenient recommendations from a social planner provide “cover” for judges)*
    - Impacting payoffs rather than just risk predictions

# Preview of Results

1. Basic fact: **Recommendations matter**
    - Algorithmic recommendations impact decisions
      *(a lenient recommendation increases lenient bail for marginal cases by 50+%)*
    - Recommendations have independent effects from algorithms themselves
2. Testing mechanisms: **Recommendations can change private costs of errors**
    - Making mistakes is less costly when decision consistent with recommendation
      *(lenient recommendations from a social planner provide "cover" for judges)*
    - Impacting payoffs rather than just risk predictions
3. Heterogeneity: **Recommendations may not impact all groups equally**
    - Judges deviate from lenient recommendation more for Black defendants than for white defendants with identical algorithmic scores

# Roadmap

1. Empirical setting: Kentucky bail decisions
2. What are the effects of algorithmic recommendations?
3. What is the mechanism behind the effect?
4. Heterogeneous effects of recommendations and implications for racial inequality

# Empirical Setting: Kentucky Bail Decisions

# Criminal justice bail decisions and algorithms

**Judge objective:** minimize pretrial detention, minimize pretrial misconduct

**Lever:** setting money bail (requires defendant to post money for release from jail)

# Criminal justice bail decisions and algorithms

**Judge objective:** minimize pretrial detention, minimize pretrial misconduct

**Lever:** setting money bail (requires defendant to post money for release from jail)

| STATE | TYPE/SCOPE OF USE |
|---|---|
| Alabama | VPRAI / Jefferson County |
| Alaska | State Created / Statewide |
| Arizona | PSA / Statewide \| VPRAI / 2 County Superior Courts |
| Arkansas | State Created / Statewide |
| California (Sample risk assessment documents from San Francisco, and Napa County) | PSA / 3 counties \| PRRS II / 2 Counties |
| Colorado (sample risk assessment documents) | CPAT / Statewide \| ODARA for DV / Statewide |
| Connecticut | State created / Statewide |
| Delaware | State created (DELPAT) / Statewide |
| District of Columbia | Developed with Urban Institute and Maxarth |
| Florida | PSA / Volusia County \| COMPAS - Sentencing / Statewide \| State Created FPRAI Being piloted / 6 Counties |
| Georgia | State created / Some counties |
| Hawaii | PSA / Statewide \| ORAS-PAT / Statewide |
| Idaho (see FOI documents below) | State created / Statewide \| Ada County / Revised IPRAI |
| Illinois | PSA / 4 counties \| VPRAI/RVRA / Most Courts |
| Indiana (sample risk assessment documents) | Mandatory use of IRAS and IYAS / Statewide |
| Iowa | PSA / 4 Counties via Pilot Program \| IRR |
| Kansas | State created / Johnson County |
| Kentucky | PSA / Statewide |
| Louisiana | PSA / New Orleans |
| Maine | ODARA (sex offenders) / Statewide \| 2019 Task Force for expansion |
| Maryland | State created / Most counties |
| Massachusetts | COMPAS / Statewide LS/CMI / Statewide |
| Michigan | COMPAS for Sentencing / Statewide |
| Minnesota (see Pretrial Release Evaluation Form and Bench Card) | MNPAT / Statewide |
| Mississippi | CRJ (Crime Justice Institute) / Statewide |
| Missouri | PSA / 1 County \| Statewide / State created \| Separate statewide system for Juvenile and Sex Offenders \| Use Oregon Public Safety Checklist for Sentencing |
| Montana | PSA / 2 Counties and 5 Pilot Counties |
| Nebraska | STRONG-R |
| Nevada | State created / Statewide Mar. 2019 by NV Supreme Court |
| New Hampshire | Yes |
| New Jersey | PSA / Statewide |
| New Mexico | PSA / 4 Counties \| ODARA for DV |
| New York | (NYC) City Created / Citywide \|State Created / State-wide for Parole |
| North Carolina | PSA / 1 County \| Developing another statewide one |
| Ohio | PSA / 3 Counties \| ORAS-PAT / Statewide |
| Oklahoma | ORAS for Pretrial Services Program + LSI/R / Statewide |
| Oregon (sample assessments) | Public Safety Checklist |
| Pennsylvania | PSA / Allegheny County \| State created / 1 County |
| Rhode Island | PSA / Statewide |
| South Carolina | State Created - Cash Bail Use |
| South Dakota | PSA / 2 Counties |
| Tennessee | PSA / 2 Counties \| State Created / One Judicial District Test |
| Texas (sample assessments) | PSA / Harris + Dallas County \| PRAISTX (derivative of ORAS) / Statewide Parole Board |
| Utah | PSA / Statewide |
| Vermont | ORAS |
| Virginia | VPRAI revised by Luminosity / Statewide \| Use Oregon Public Safety Checklist for Sentencing |
| Washington | PSA / 3 Counties |
| West Virginia | LS/CMI |
| Wisconsin (See sample assessment documents) | PSA / 4 Counties \| COMPAS / Statewide |
| Wyoming | COMPAS for Prisoners / Statewide |
| Federal | PTRA |

*Source: Epic (2020)*

# Criminal justice bail decisions and algorithms

**Judge objective:** minimize pretrial detention, minimize pretrial misconduct

**Lever:** setting money bail (requires defendant to post money for release from jail)

| STATE | TYPE/SCOPE OF USE |
|---|---|
| Alabama | VPRAI / Jefferson County |
| Alaska | State Created / Statewide |
| Arizona | PSA / Statewide \| VPRAI / 2 County Superior Courts |
| and Napa County) | |
| Colorado (sample risk assessment documents) | CPAT / Statewide \| ODARA for DV / Statewide |
| Connecticut | State created / Statewide |
| Delaware | State created (DELPAT) / Statewide |
| District of Columbia | Developed with Urban Institute and Maxarth |
| Florida | PSA / Volusia County \| COMPAS - Sentencing / Statewide \| State Created FPRAI Being piloted / 6 Counties |

| STATE | TYPE/SCOPE OF USE |
|---|---|
| Georgia | State created / Some counties |
| Hawaii | PSA / Statewide \| ORAS-PAT / Statewide |
| Idaho (see FOI documents below) | State created / Statewide \| Ada County / Revised IPRAI |
| Illinois | PSA / 4 counties \| VPRAI/RVRA / Most Courts |
| Indiana (sample risk | Mandatory use of IRAS and |
| Louisiana | PSA / New Orleans |
| Maine | ODARA (sex offenders) / Statewide \| 2019 Task Force for expansion |
| Maryland | State created / Most counties |
| Massachusetts | COMPAS / Statewide LS/CMI / Statewide |
| Michigan | COMPAS for Sentencing / Statewide |
| Minnesota (see Pretrial Release Evaluation Form and Bench Card) | MNPAT / Statewide |
| Mississippi | CRJ (Crime Justice Institute) / Statewide |
| Missouri | PSA / 1 County \| Statewide / State created \| Separate statewide system for Juvenile and Sex Offenders \| Use Oregon Public Safety Checklist for Sentencing |

| | |
|---|---|
| Montana | PSA / 2 Counties and 5 Pilot Counties |
| Nebraska | STRONG-R |
| Nevada | State created / Statewide Mar. 2019 by NV Supreme Court |
| New Hampshire | Yes |
| New Jersey | PSA / Statewide |
| | another statewide one |
| Ohio | PSA / 3 Counties \| ORAS-PAT / Statewide |
| Oklahoma | ORAS for Pretrial Services Program + LSI/R / Statewide |
| Oregon (sample assessments) | Public Safety Checklist |
| Pennsylvania | PSA / Allegheny County \| State created / 1 County |
| Rhode Island | PSA / Statewide |
| South Carolina | State Created - Cash Bail Use |
| South Dakota | PSA / 2 Counties |
| Tennessee | PSA / 2 Counties \| State Created / One Judicial District Test |
| Texas (sample assessments) | PSA / Harris + Dallas County \| PRAISTX (derivative of ORAS) / Statewide Parole Board |
| Utah | PSA / Statewide |
| Vermont | ORAS |

| | |
|---|---|
| Virginia | VPRAI revised by Luminosity / Statewide \| Use Oregon Public Safety Checklist for Sentencing |
| Washington | PSA / 3 Counties |
| West Virginia | LS/CMI |
| Wisconsin (See sample assessment documents) | PSA / 4 Counties \| COMPAS / Statewide |
| Wyoming | COMPAS for Prisoners / Statewide |
| Federal | PTRA |

Common risk score goal: "data-driven way to advance pretrial release"

*Source: Epic (2020)*

# Criminal justice bail decisions and algorithms

**Judge objective:** minimize pretrial detention, minimize pretrial misconduct

**Lever:** setting money bail (requires defendant to post money for release from jail)

Common risk score goal: "data-driven way to advance pretrial release"

Predict misconduct based on observable data

| STATE | TYPE/SCOPE OF USE |
|---|---|
| Alabama | VPRAI / Jefferson County |
| Alaska | State Created / Statewide |
| Arizona | PSA / Statewide \| VPRAI / 2 County Superior Courts |
| and Napa County) | |
| Colorado (sample risk assessment documents) | CPAT / Stat… DV / Statewi… |
| Connecticut | State create… |
| Delaware | State create… / Statewide |
| District of Columbia | Developed with Urban Institute and Maxarth |
| Florida | PSA / Volusia County \| COMPAS - Sentencing / Statewide \| State Created FPRAI Being piloted / 6 Counties |

| STATE | TYPE/SCOPE OF USE |
|---|---|
| Georgia | State created / Some counties |
| Hawaii | PSA / Statewide \| ORAS-PAT / Statewide |
| Idaho (see FOI documents below) | State created / Statewide \| Ada County / Revised IPRAI |
| Illinois | PSA / 4 counties \| VPRAI/RVRA / Most Courts |
| Maine | ODARA (sex offenders) / Statewide \| 2019 Task Force for expansion |
| Release Evaluation Form and Bench Card) | MNPAT / Statewide |
| Mississippi | CRJ (Crime Justice Institute) / Statewide |
| Missouri | PSA / 1 County \| Statewide / State created \| Separate statewide system for Juvenile and Sex Offenders \| Use Oregon Public Safety Checklist for Sentencing |

| | |
|---|---|
| Montana | PSA / 2 Counties and 5 Pilot Counties |
| Nebraska | STRONG-R |
| Nevada | State created / Statewide Mar. 2019 by NV Supreme Court |
| New Hampshire | Yes |
| New Jersey | PSA / Statewide |
| Ohio | PSA / 3 Counties \| ORAS-PAT / Statewide |
| Rhode Island | PSA / Statewide |
| South Carolina | State Created - Cash Bail Use |
| South Dakota | PSA / 2 Counties |
| Tennessee | PSA / 2 Counties \| State Created / One Judicial District Test |
| Texas (sample assessments) | PSA / Harris + Dallas County \| PRAISTX (derivative of ORAS) / Statewide Parole Board |
| Utah | PSA / Statewide |
| Vermont | ORAS |

| | |
|---|---|
| Virginia | VPRAI revised by Luminosity / Statewide \| Use Oregon Public Safety Checklist for Sentencing |
| | PSA / 3 Counties |
| | LS/CMI |
| …ee sample | PSA / 4 Counties \| COMPAS / Statewide |
| Wyoming | COMPAS for Prisoners / Statewide |
| Federal | PTRA |

*Source: Epic (2020)*

# The Kentucky Algorithm

After person booked, pretrial services officer calculates a risk score

***Kentucky Pretrial Risk Assessment tool***
*(March 2011-May 2013)*

# The Kentucky Algorithm

After person booked, pretrial services officer calculates a risk score

***Kentucky Pretrial Risk Assessment tool***
*(March 2011-May 2013)*

- Not complex black-box ML tool – it is a "checklist tool" (or "rule-based formula")
- Total points and convert to levels:
  - *0-5: low*
  - *6-13: moderate*
  - *14-24: high*
- Scores have relative, not absolute meaning *(e.g., high is riskier than low)*
- **Only levels shared with judges**

| **Risk Component** | **Points** |
|---|---|
| No verified address | 2 |
| No verified means of support | 1 |
| ABC Felony charge | 1 |
| Pending case | 7 |
| Prior/active mis/felony FTA | 2 |
| Prior FTA traffic violation | 1 |
| Prior misdemeanors | 2 |
| Prior felonies | 1 |
| Prior violent convictions | 1 |
| History of drug/alcohol abuse | 2 |
| Prior felony escape conviction | 3 |
| On probation/parole | 1 |

# House Bill introduces recommendation for some cases

Before June 2011:

- Judge receives info about defendant, incident, risk level and makes a bail decision in a few minutes
    - *Risk level: Kentucky Pretrial Risk Assessment tool*
    - Judge decides whether to set money bail

# House Bill introduces recommendation for some cases

Before June 2011:

- Judge receives info about defendant, incident, risk level and makes a bail decision in a few minutes
    - *Risk level: Kentucky Pretrial Risk Assessment tool*
    - Judge decides whether to set money bail

Starting June 2011:

- House Bill *(legislature action)* **recommends no money bail ("lenient bail") for low and moderate risk level cases**
    - Judges could deviate by saying a few words *(no large admin cost)*
    - No recommendation for high risk cases

# Variation in recommendation over time and scores

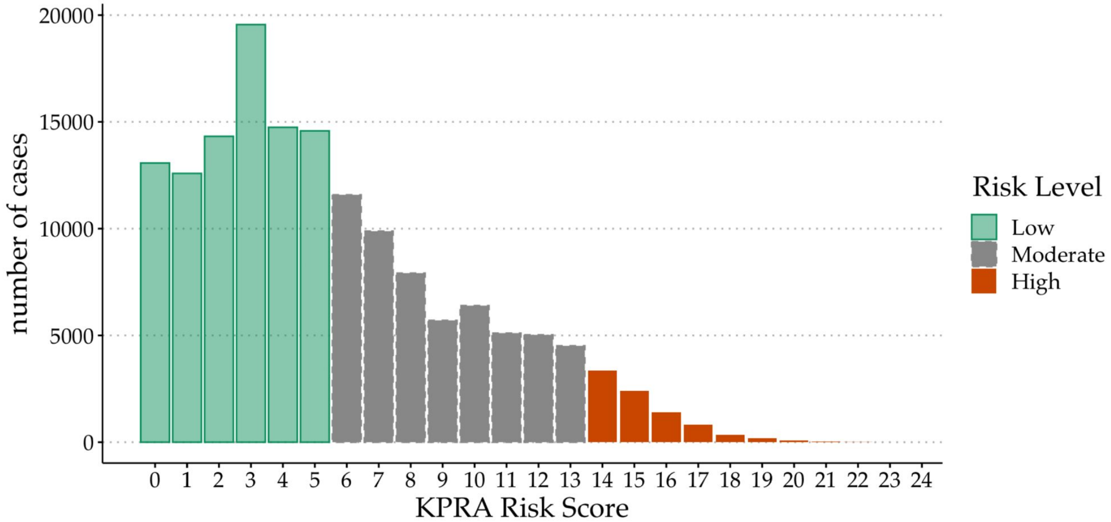
No recommendation

June 2011

Recommendation
*(lenient bail for low/moderate cases)*

# Variation in recommendation over time and scores

No recommendation | June 2011 | Recommendation *(lenient bail for low/moderate cases)*

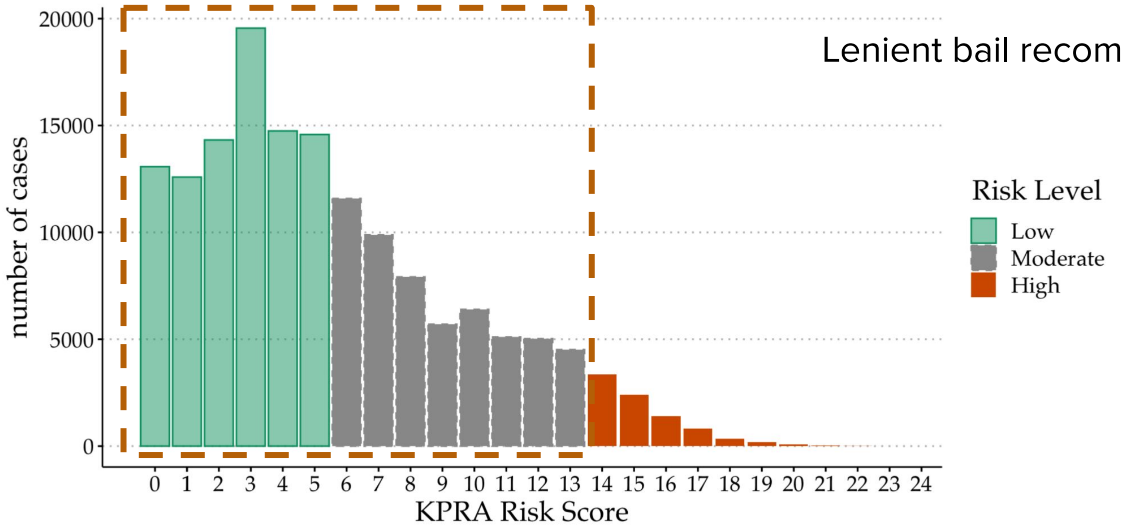


# Variation in recommendation over time and scores

No recommendation

June 2011

Recommendation
*(lenient bail for low/moderate cases)*



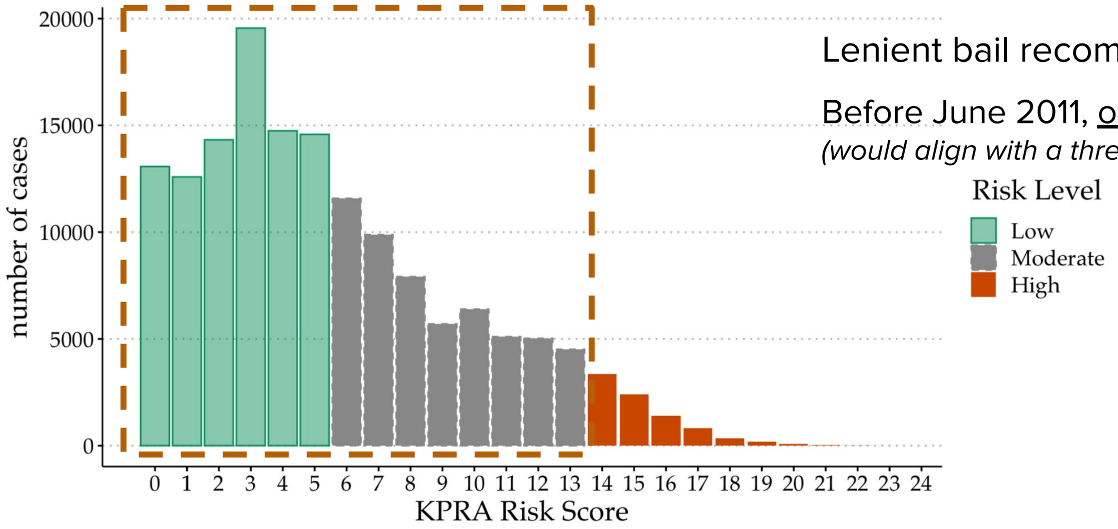

Lenient bail recommended for **90% of cases**

# Variation in recommendation over time and scores

No recommendation

June 2011

Recommendation
*(lenient bail for low/moderate cases)*




Lenient bail recommended for **<u>90% of cases</u>**

Before June 2011, <u>only **32%**</u> got lenient bail
*(would align with a threshold of score<4)*

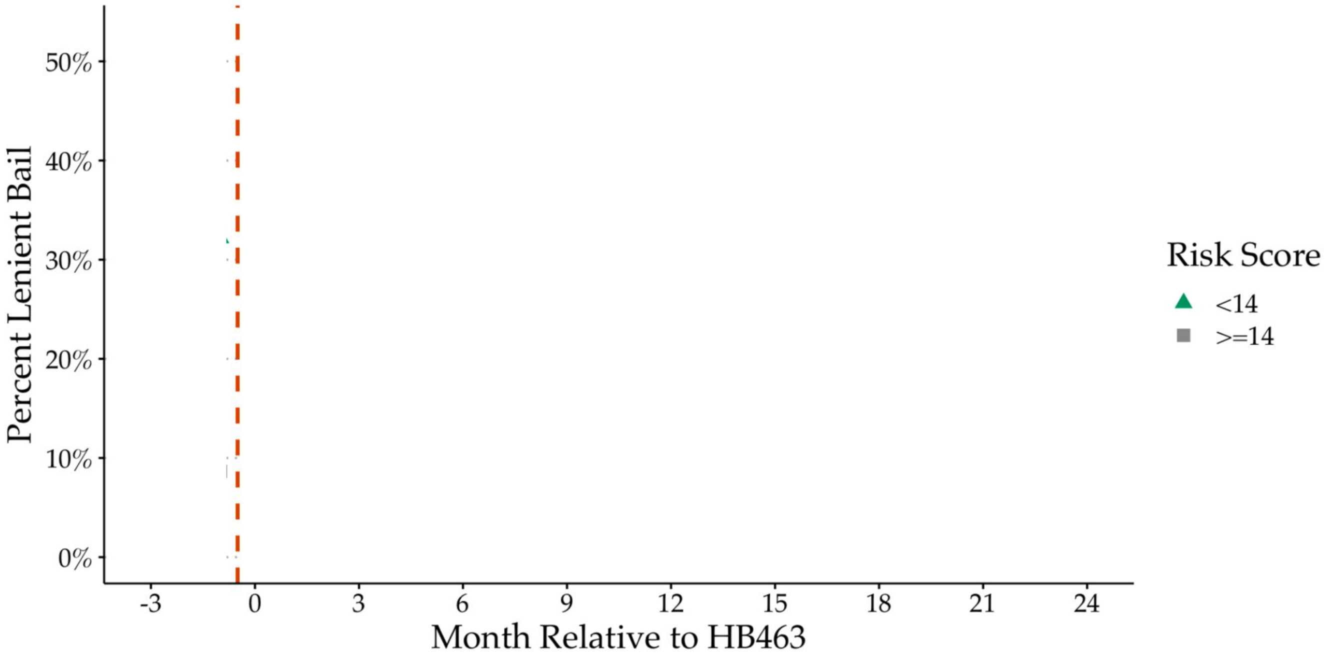
# What are the effects of recommendations?

# Difference-in-differences approach

- Cases with scores < 14 get a lenient recommendation
- Cases with scores >= 14 do not

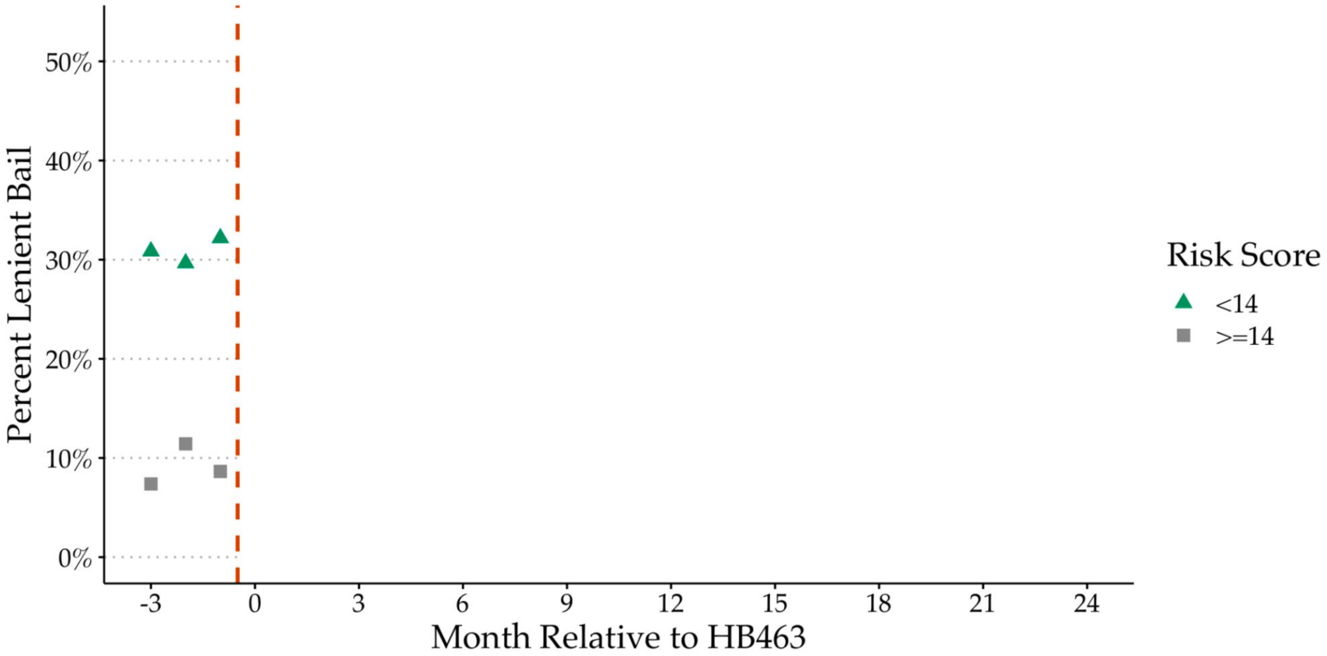
# Difference-in-differences approach

- Cases with scores < 14 get a lenient recommendation
- Cases with scores >= 14 do not



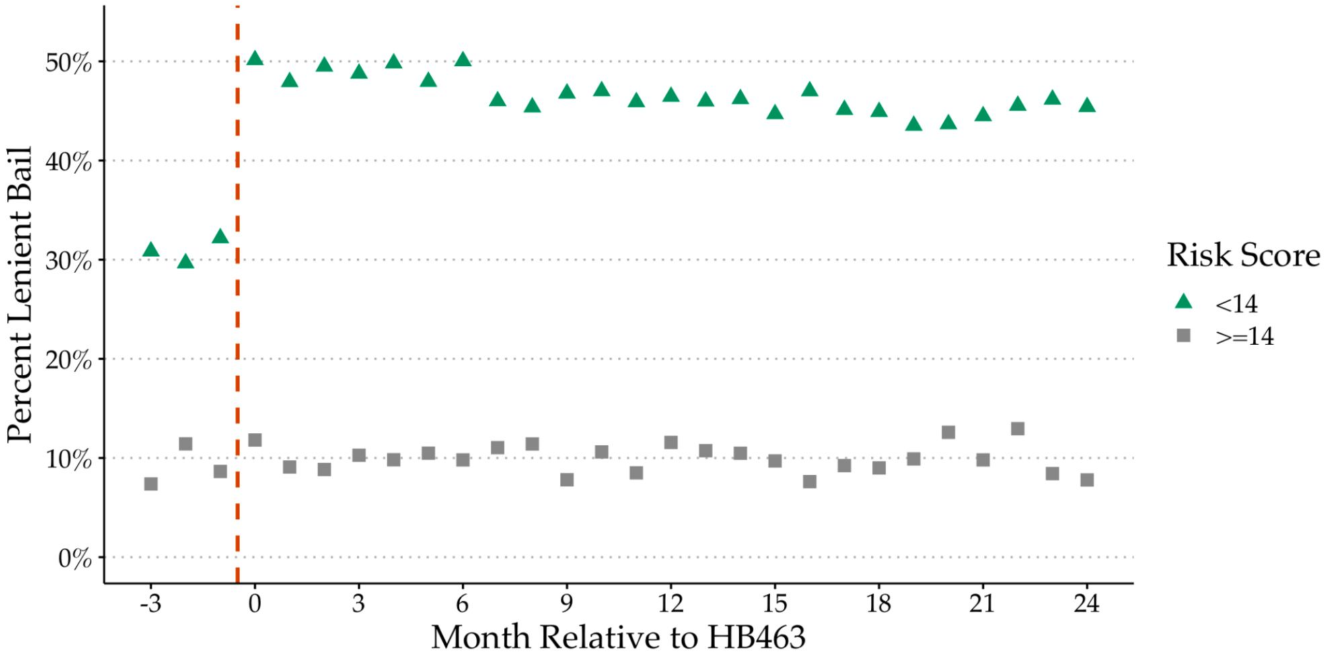
# Difference-in-differences approach

- Cases with scores < 14 get a lenient recommendation
- Cases with scores >= 14 do not



# Difference-in-differences approach

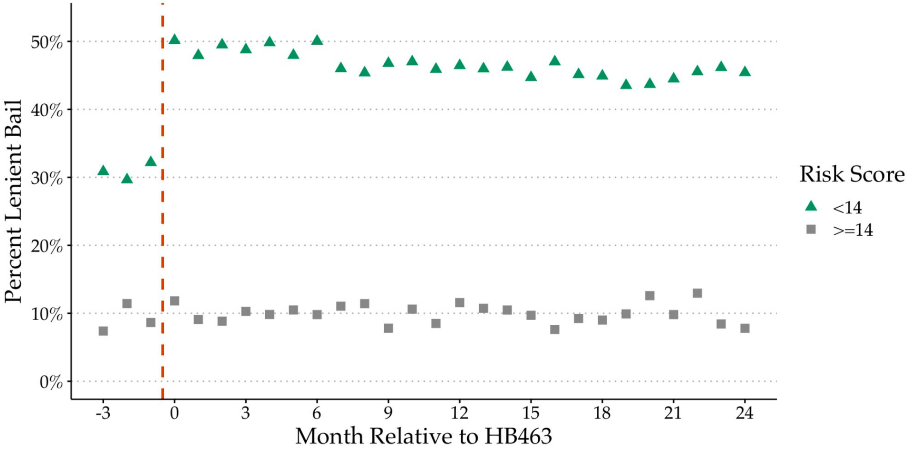
- Cases with scores < 14 get a lenient recommendation
- Cases with scores >= 14 do not

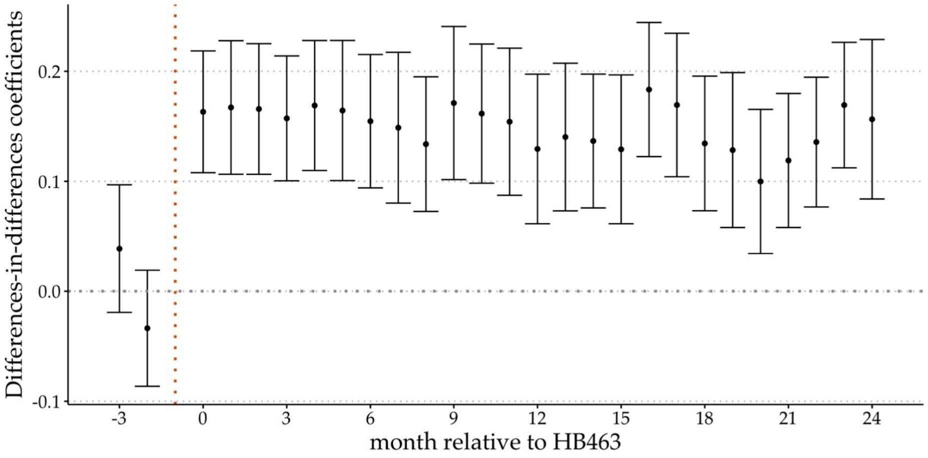


# Difference-in-differences approach

- Cases with scores < 14 get a lenient recommendation
- Cases with scores >= 14 do not

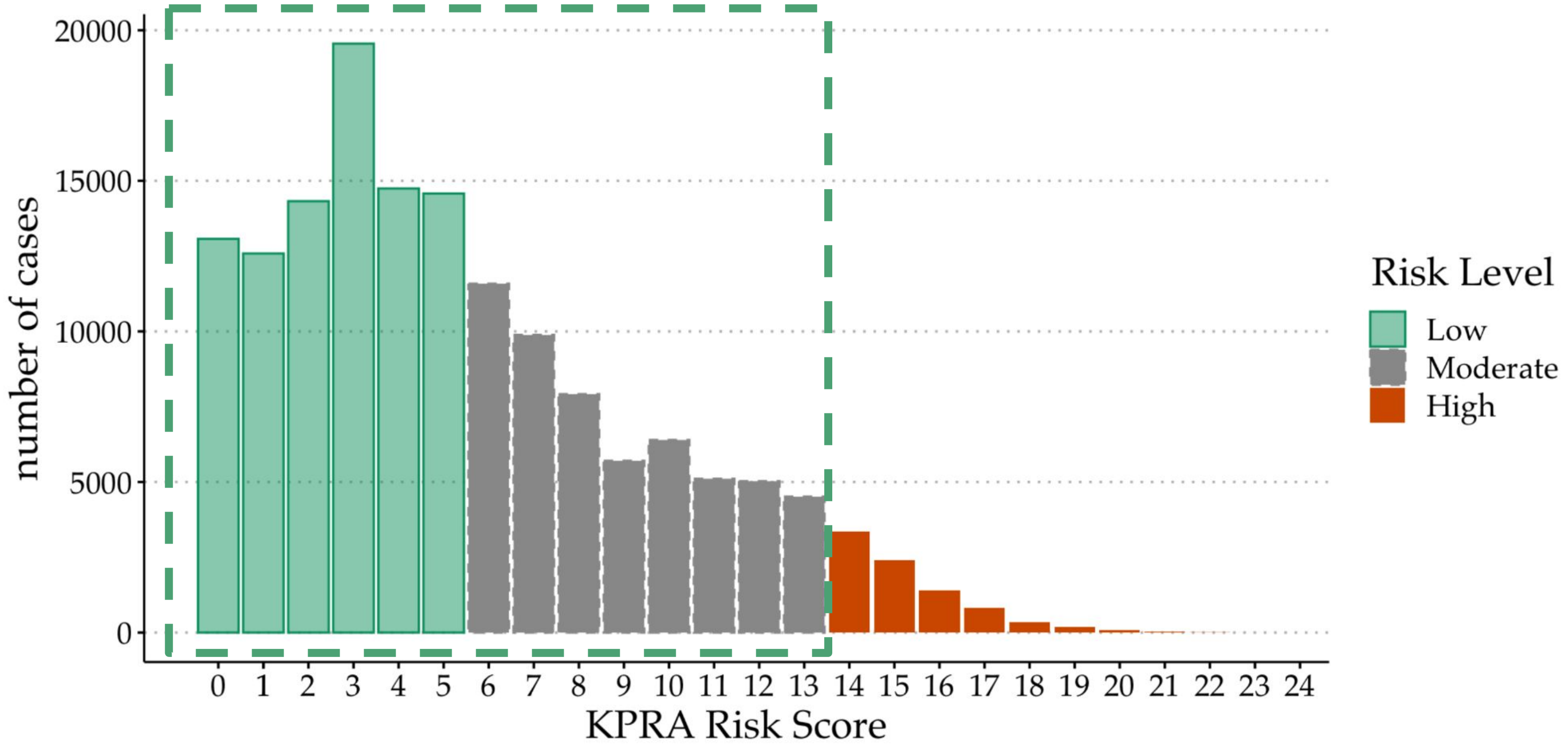
$$lenient_{itj} = \sum_{m \neq -1} [\beta_m \times I(score_i < 14)] + X_{itj} + \epsilon_{itj}$$



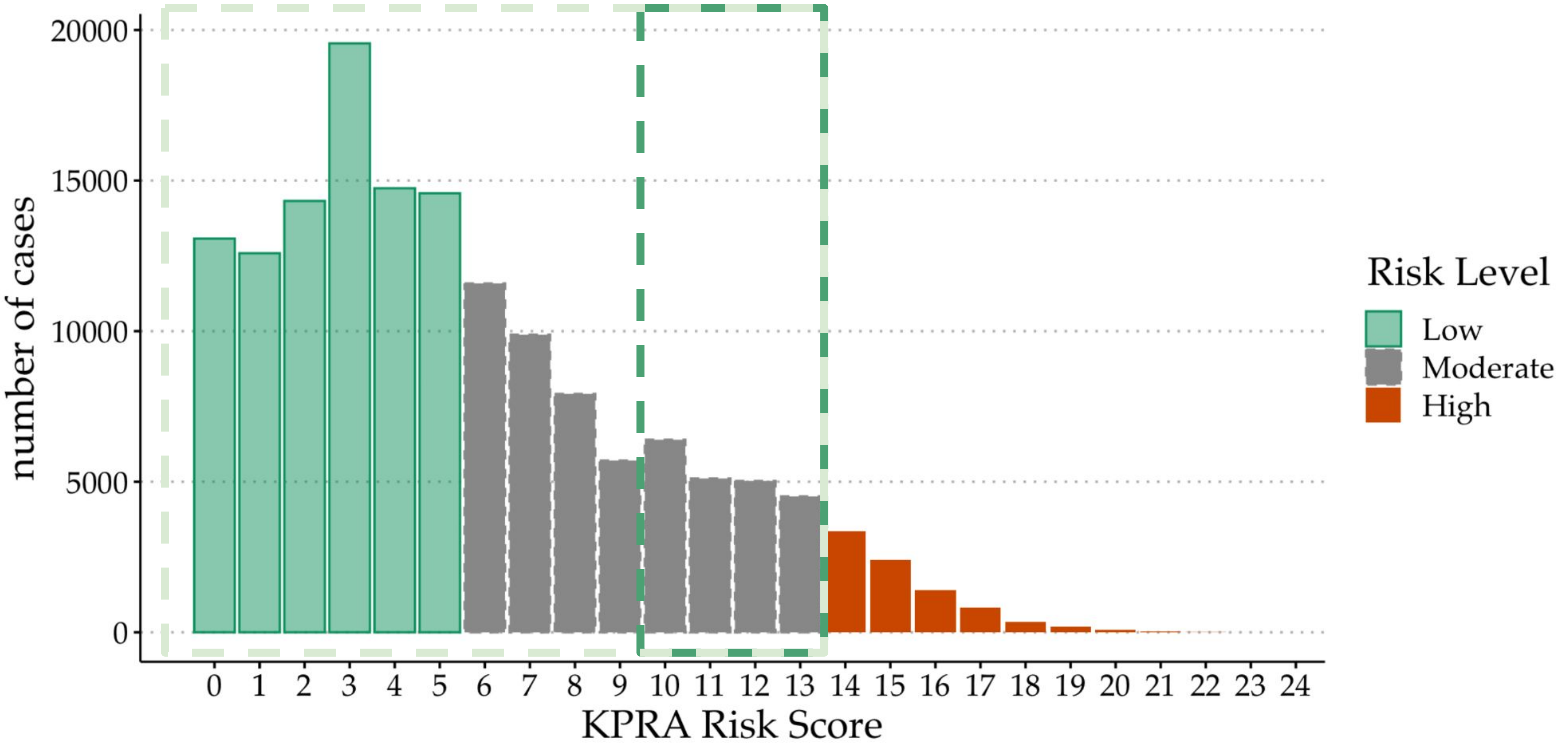




<u>Pooled DD:</u> **15 pp increase / 50% increase** (off the 30% baseline)

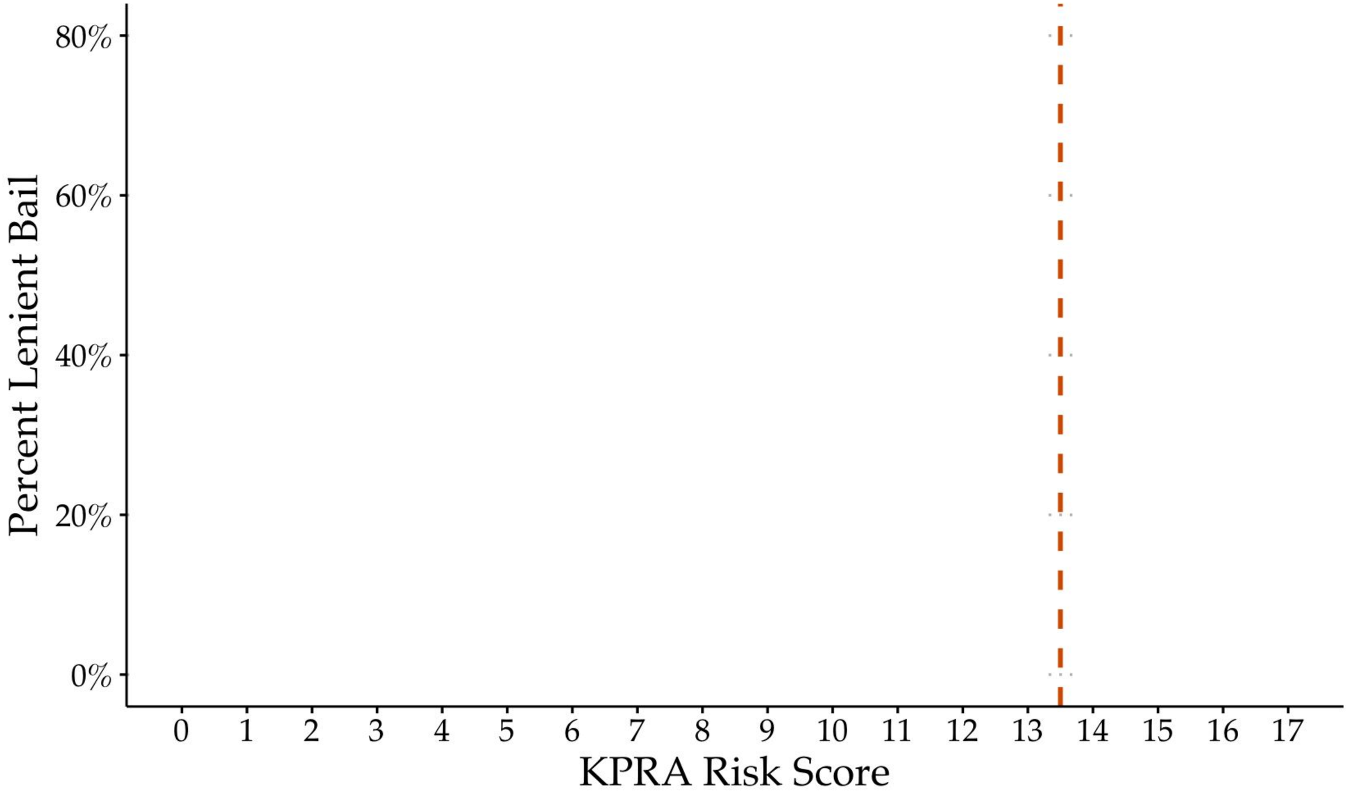
# Heterogeneity in effects across the risk score distribution



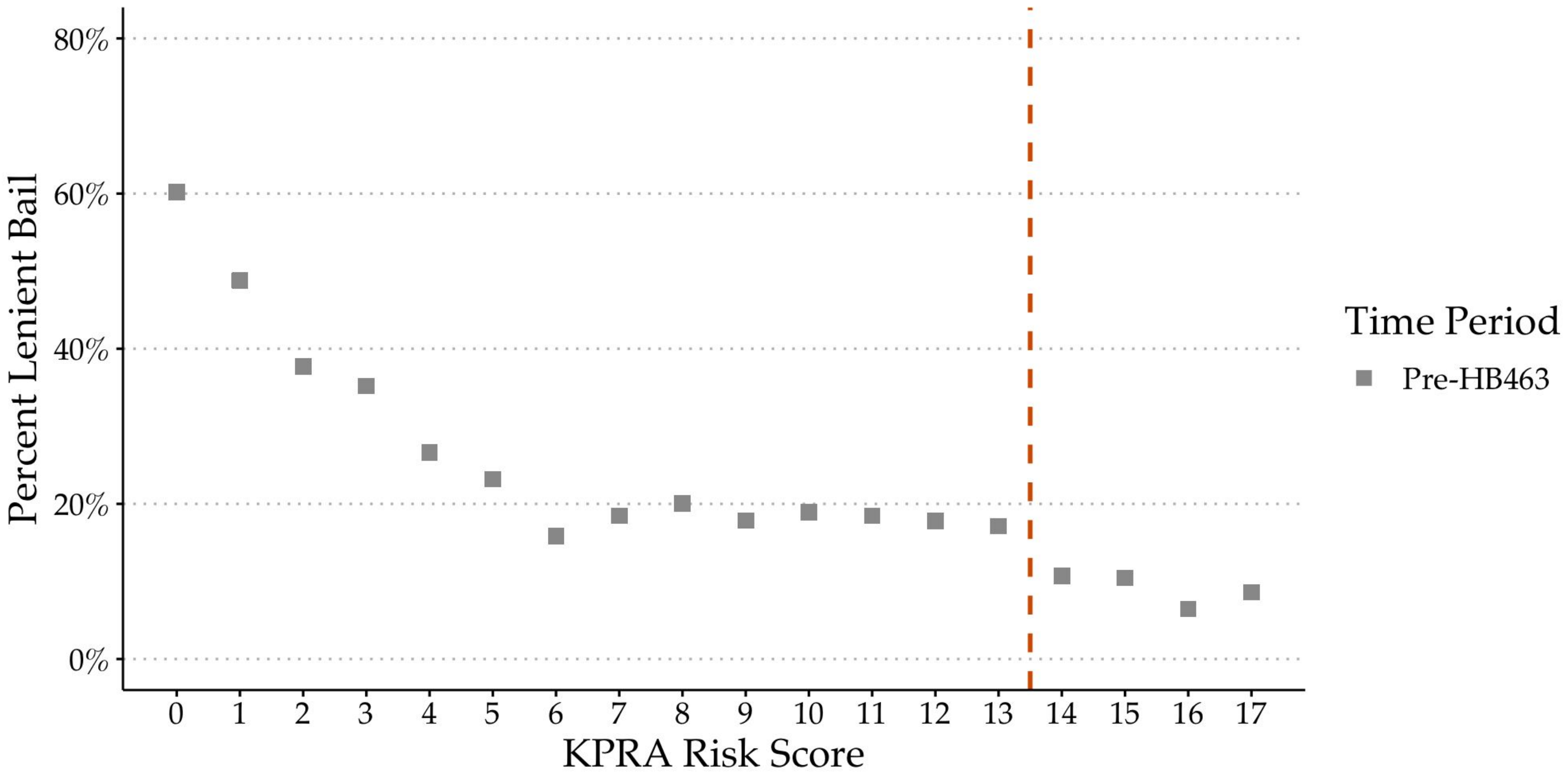
# Heterogeneity in effects across the risk score distribution



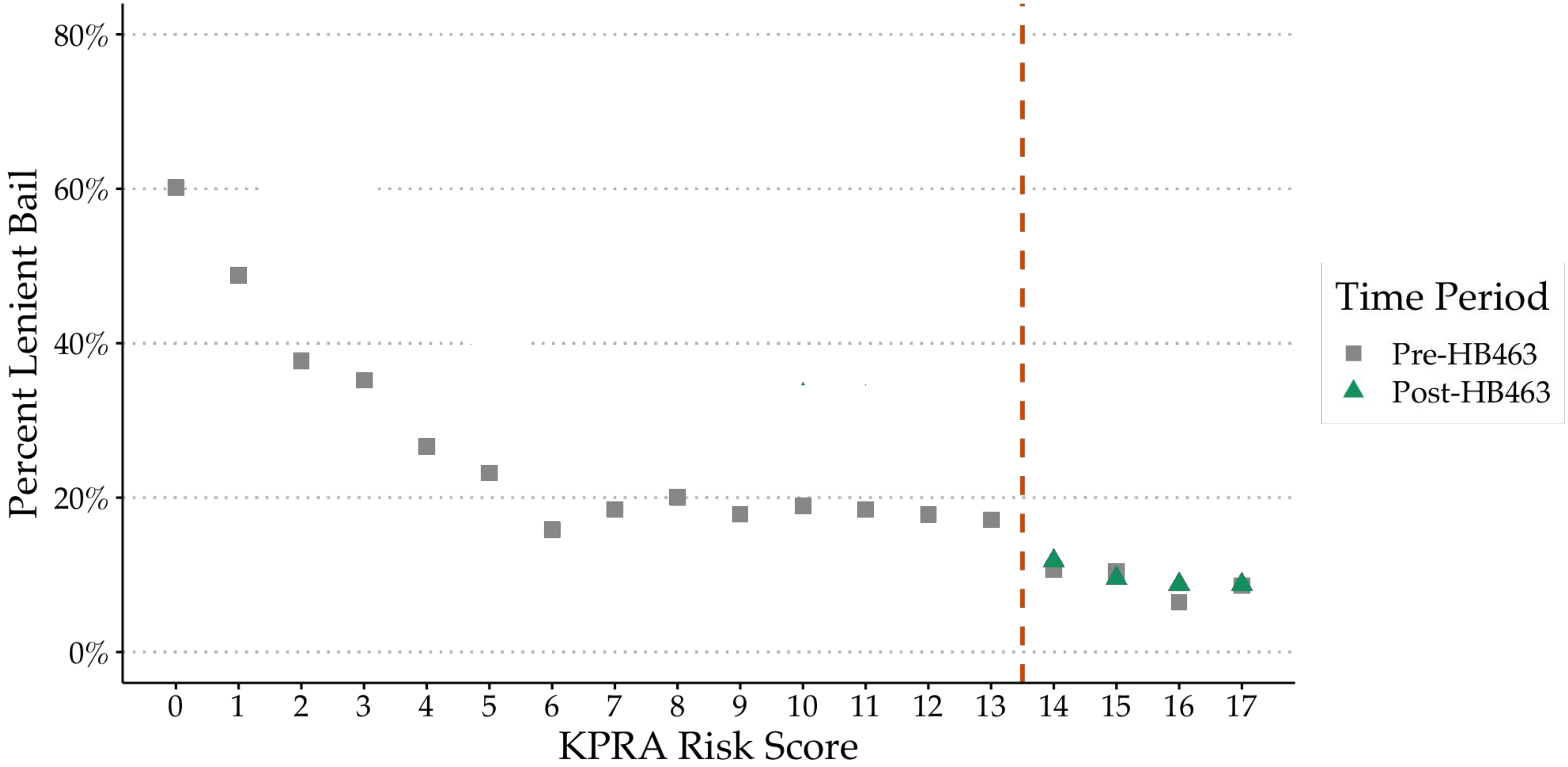
# Heterogeneity in effects across the risk score distribution



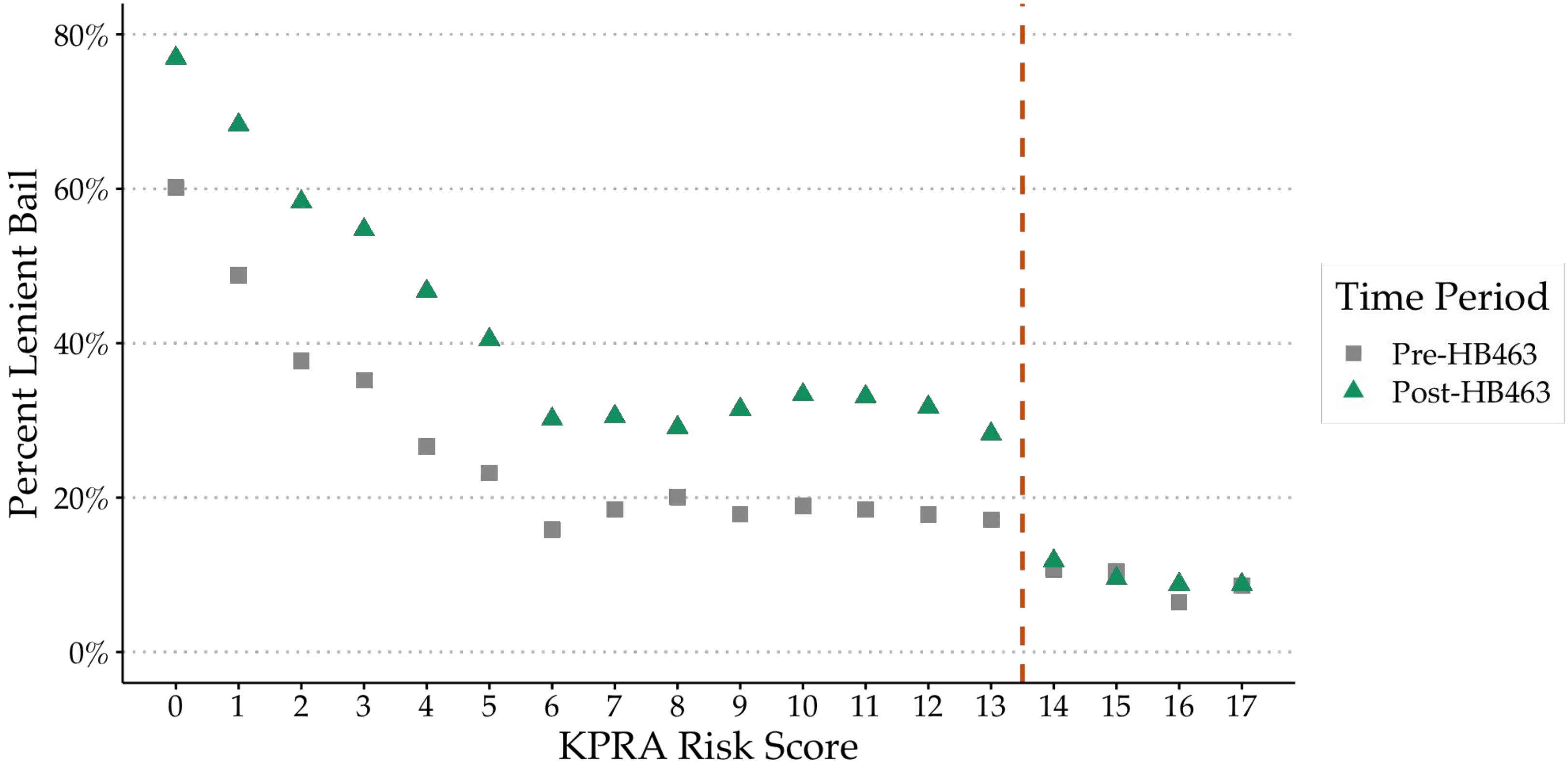
# Heterogeneity in effects across the risk score distribution



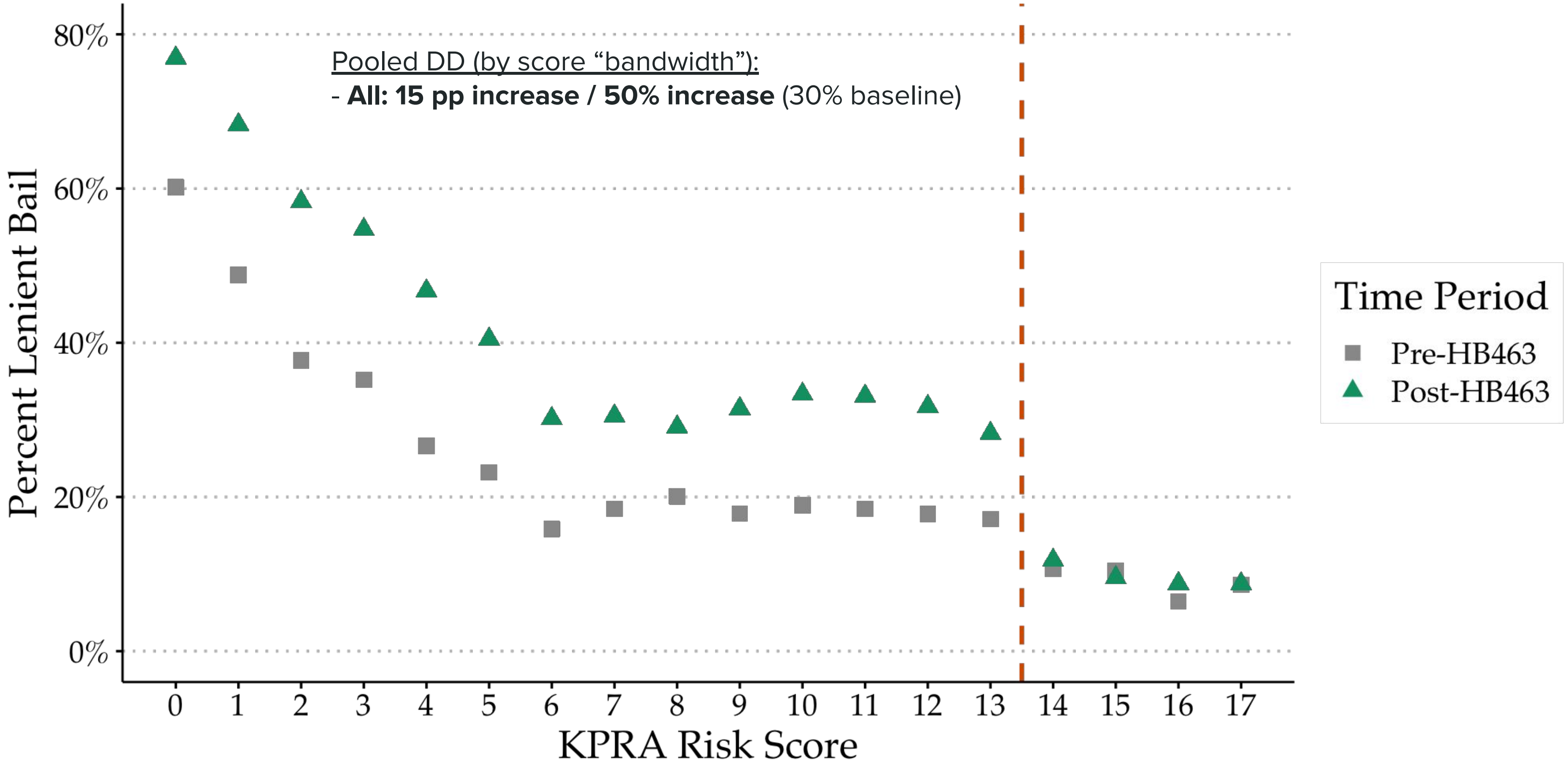
# Heterogeneity in effects across the risk score distribution



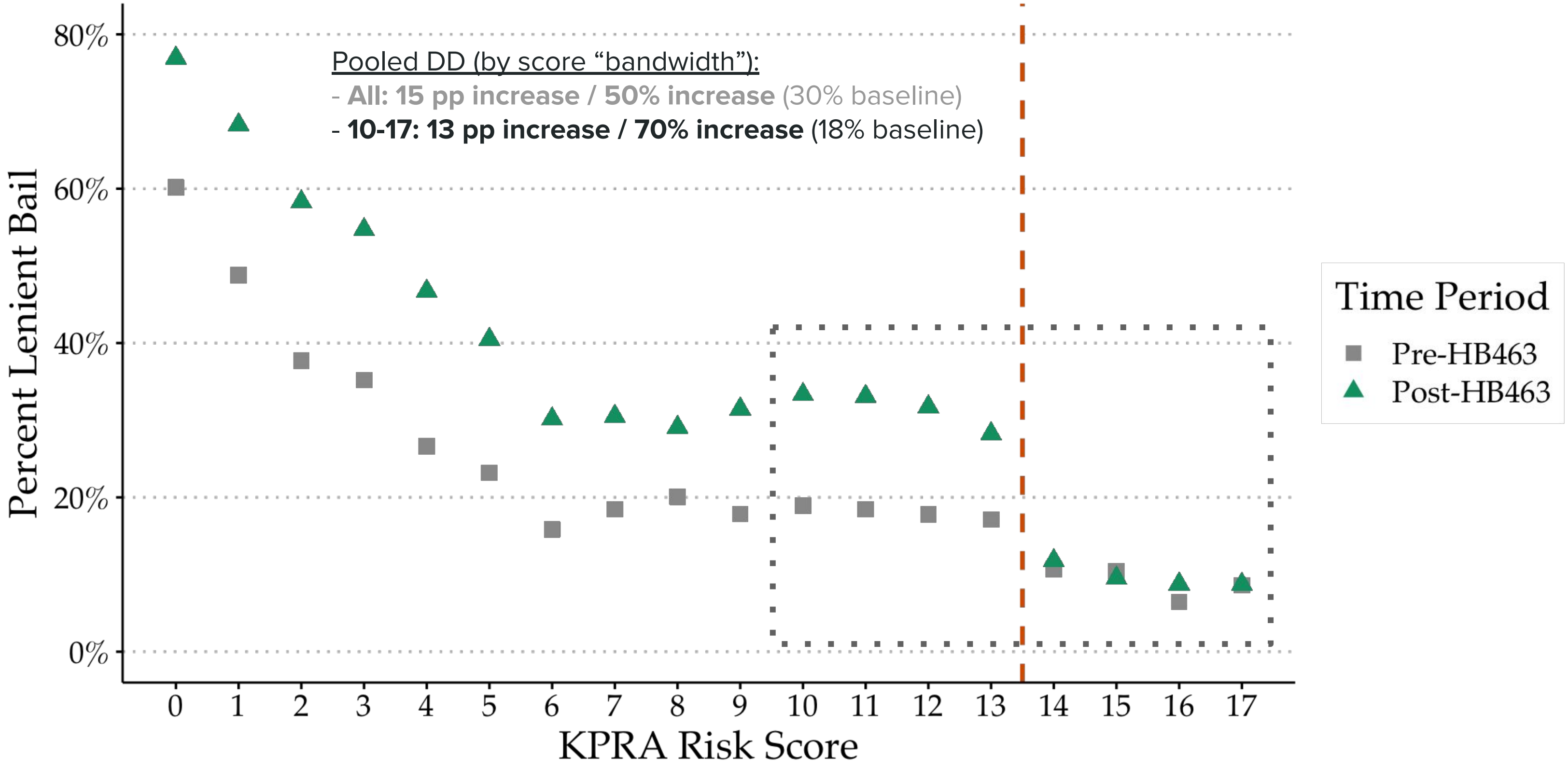
# Heterogeneity in effects across the risk score distribution



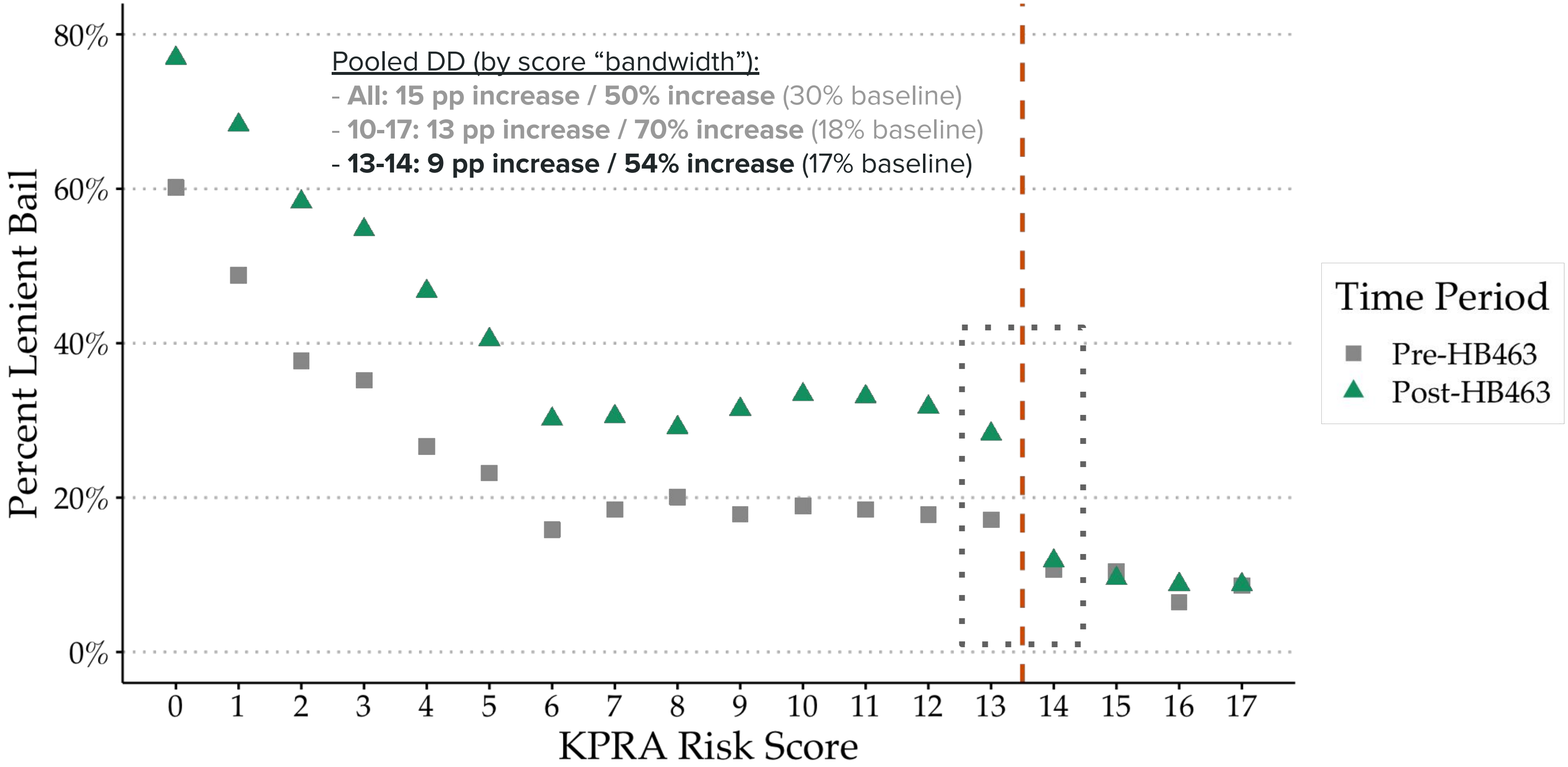
# Heterogeneity in effects across the risk score distribution



# Heterogeneity in effects across the risk score distribution



# Heterogeneity in effects across the risk score distribution



# What is the mechanism behind the effect?

# Why might recommendations *change judge decisions*?

1. Administrative cost to deviate

# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

# Why might recommendations *change judge decisions*?

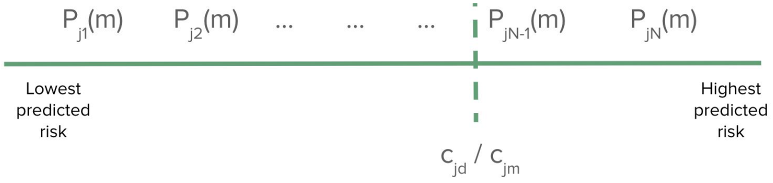
1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

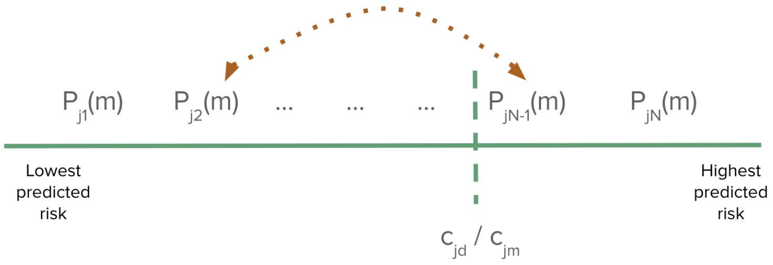
2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)



# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)




=> change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

# Why might recommendations *change judge decisions?*

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

   => change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; "type II errors")*

# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

   => change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; "type II errors")*
   - Sticking neck out more if recommended detention (additional penalty)

NEWS

**Darrell Brooks Should Not Have Been Released on Low Bail, Milwaukee DA Admits**

BY KATHERINE FUNG ON 11/22/21 AT 2:02 PM EST

*"[Bail] in this case is not consistent with … the risk assessment of the defendant prior to the setting of bail."*

# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

   => change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; "type II errors")*
   - Sticking neck out more if recommended detention (additional penalty)
   - Not as risky if recommended release (algorithm designer gives reputational cover)

*In New York City court observations,*

***"judges routinely stated that they only ordered people to be released [...] because the law forced them to."***
*(Corvert 2022)*

# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

   => change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

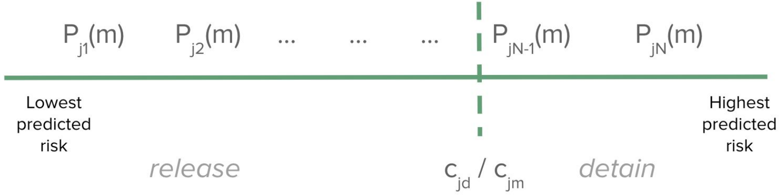
3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; “type II errors”)*
   - Sticking neck out more if recommended detention (additional penalty)
   - **Not as risky if recommended release (algorithm designer gives reputational cover)**

# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

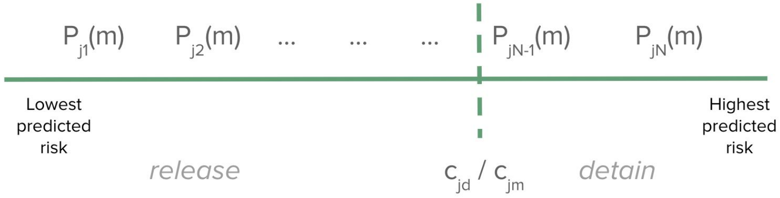
low/moderate risk




=> change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; "type II errors")*
   - Sticking neck out more if recommended detention (additional penalty)
   - **Not as risky if recommended release (algorithm designer gives reputational cover)**

high risk

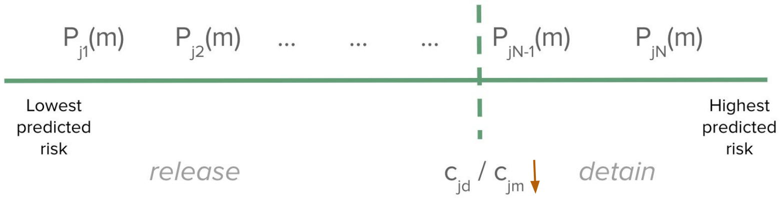


# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

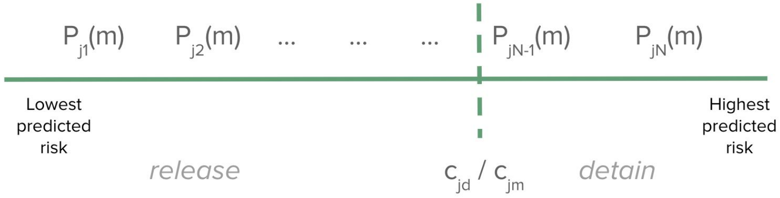
**low/moderate risk**




=> change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; "type II errors")*
    - Sticking neck out more if recommended detention (additional penalty)
    - **Not as risky if recommended release (algorithm designer gives reputational cover)**

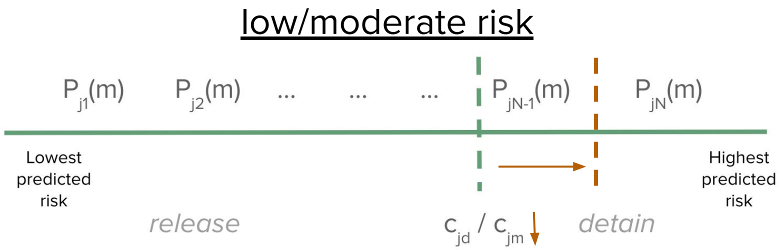
**high risk**



# Why might recommendations *change judge decisions*?

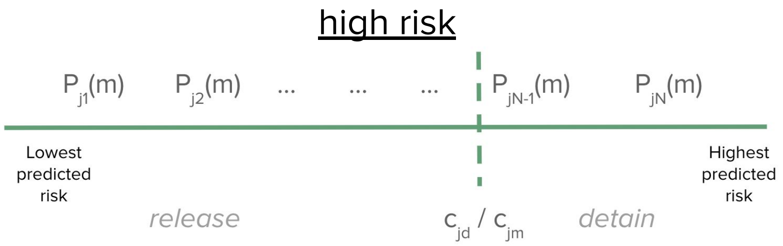
1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)




=> change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; "type II errors")*
   - Sticking neck out more if recommended detention (additional penalty)
   - **Not as risky if recommended release (algorithm designer gives reputational cover)**




=> change in *composition* of decisions, **no effect for high risk**

# Testing the dueling predictions

Want to estimate effect of recommendation introduction on high risk group *(not covered by recommendation)*

# Testing the dueling predictions

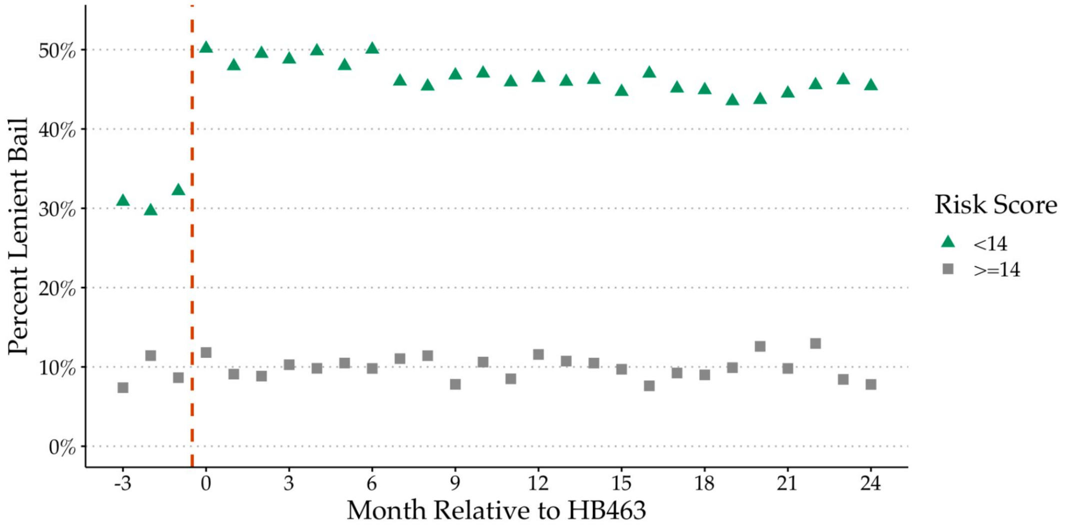
Want to estimate effect of recommendation introduction on high risk group *(not covered by recommendation)*

Methodology = **leverage variation over time (RD in time)**

# Testing the dueling predictions

Want to estimate effect of recommendation introduction on high risk group *(not covered by recommendation)*

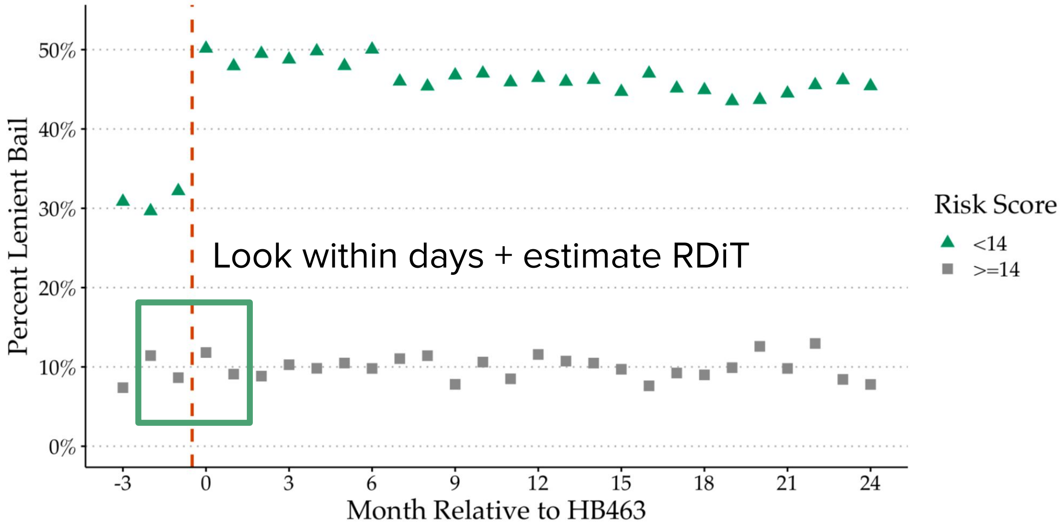
Methodology = **leverage variation over time (RD in time)**



# Testing the dueling predictions

Want to estimate effect of recommendation introduction on high risk group *(not covered by recommendation)*

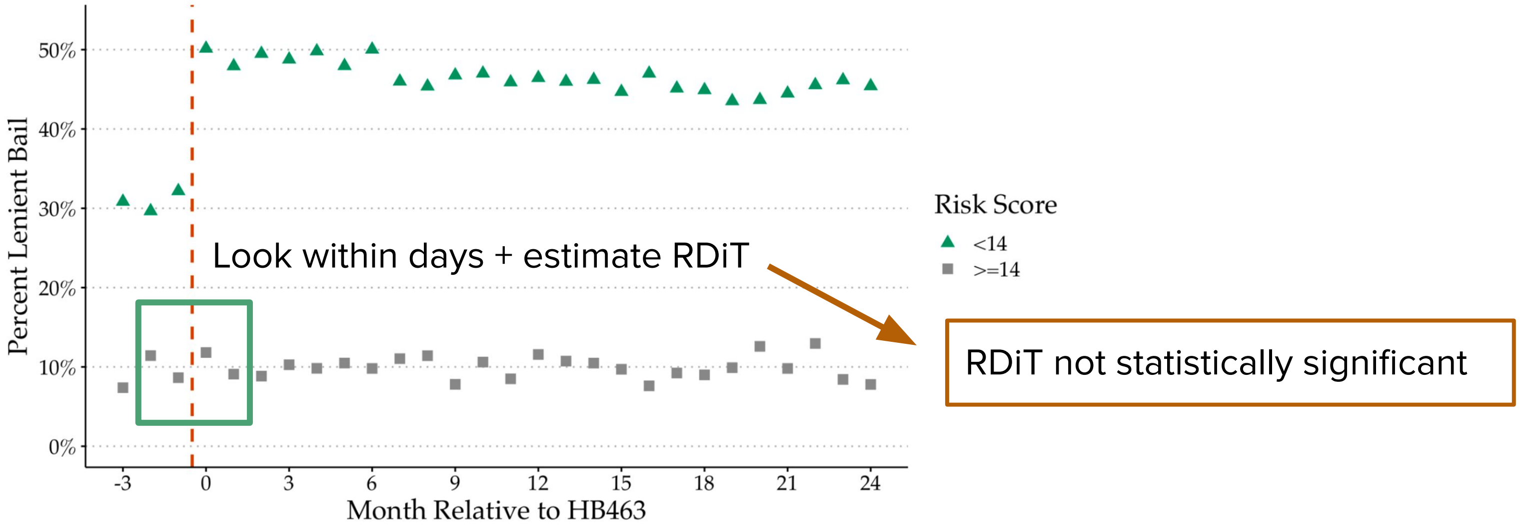
Methodology = **leverage variation over time (RD in time)**



# Testing the dueling predictions

Want to estimate effect of recommendation introduction on high risk group *(not covered by recommendation)*

Methodology = **leverage variation over time (RD in time)**



# Why might recommendations *change judge decisions*?

1. ~~Administrative cost to deviate~~

2. Recommendations increase algorithms use in setting perceived probabilities (upweight algorithm output)

   => change in *allocation* but not *composition* of decisions, **decrease in lenient bail for high risk**

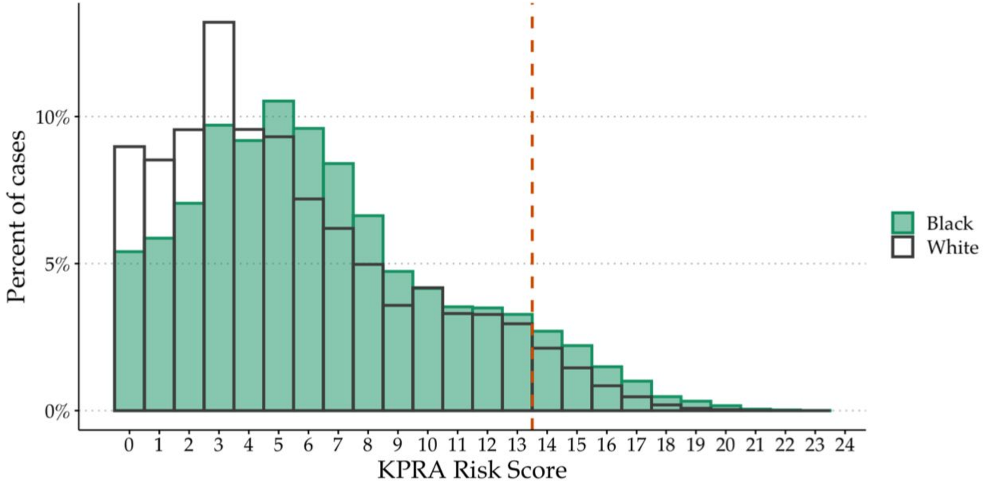
3. Recommendations change misconduct costs in the event of a bad outcome *(someone was released and commits misconduct; "type II errors")*
   - Sticking neck out more if recommended detention (additional penalty)
   - **Not as risky if recommended release (algorithm designer gives reputational cover)**

   => change in *composition* of decisions, **no effect for high risk**

# Heterogeneous recommendation effects and implications for racial inequality

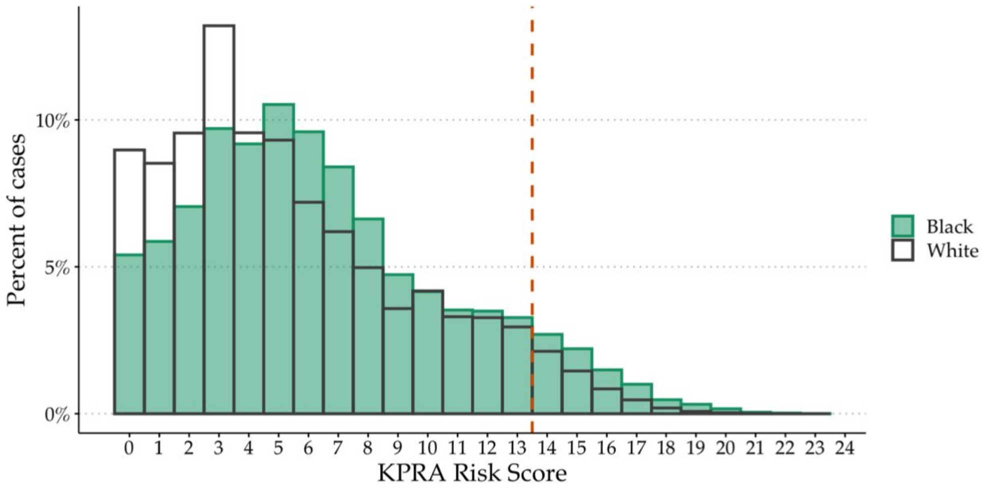
# Racial disparities in risk scores, recommendations, and outcomes

Concern about the algorithm distribution:
usage might widen racial disparities



# Racial disparities in risk scores, recommendations, and outcomes

Concern about the algorithm distribution:
usage might widen racial disparities



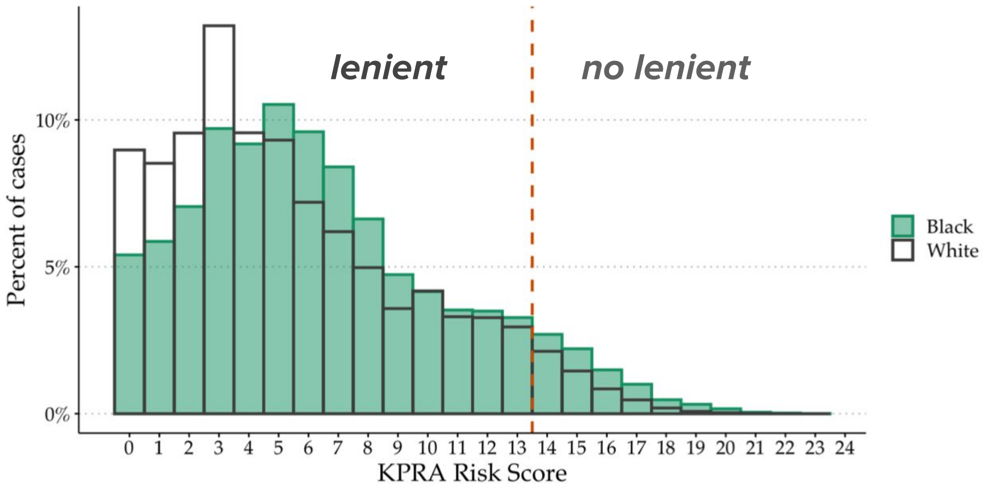

After the recommendations implemented,

Black people were **9.3 pp less likely** to get lenient bail (36.7% vs. 46%) than white people

***How much is due to different recommendations?***

# Racial disparities in risk scores, recommendations, and outcomes

Concern about the algorithm distribution:
usage might widen racial disparities




After the recommendations implemented,

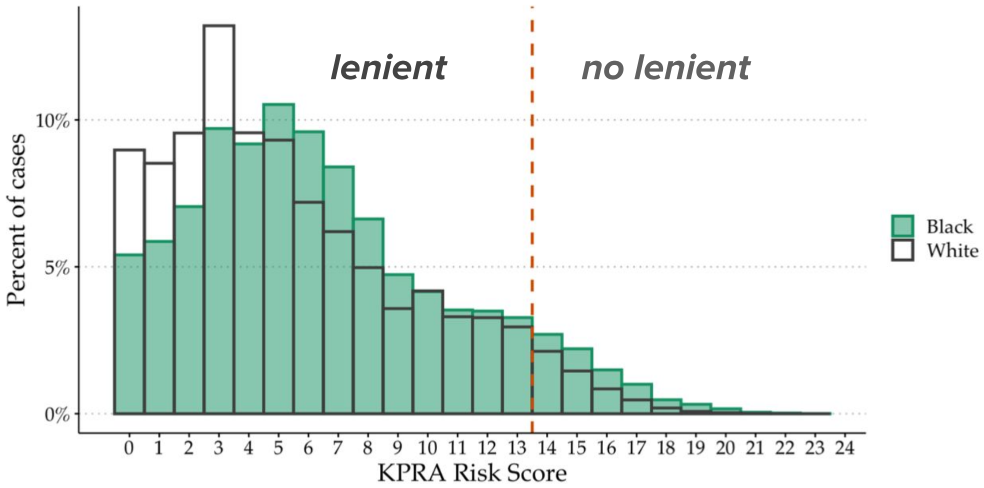
Black people were **9.3 pp less likely** to get lenient bail (36.7% vs. 46%) than white people

*If bail automatically set by recommendations (low/mod => lenient; high => no lenient),*

Black people would've been **3.3 pp less likely** to get lenient bail (91.5% vs. 94.8%) than white people

# Racial disparities in risk scores, recommendations, and outcomes

Concern about the algorithm distribution:
usage might widen racial disparities




After the recommendations implemented,

Black people were **9.3 pp less likely** to get lenient bail (36.7% vs. 46%) than white people

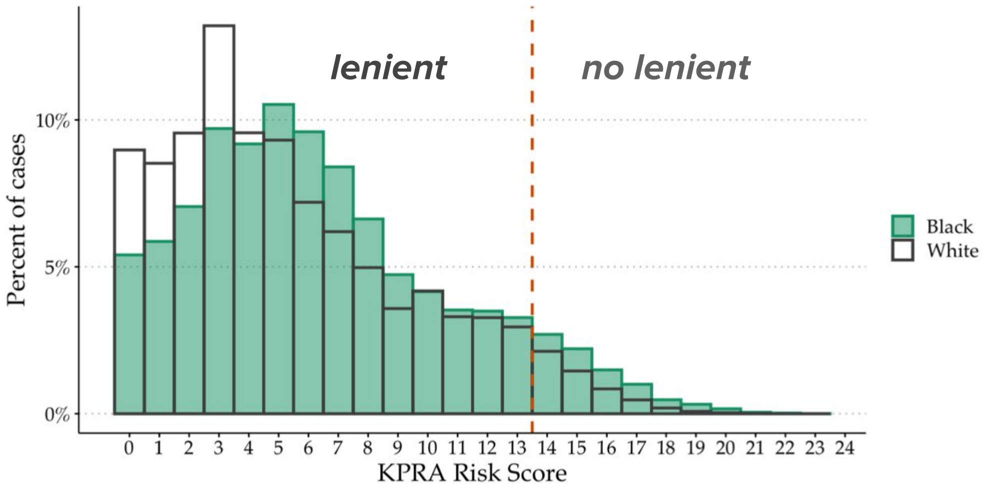
*If bail automatically set by recommendations (low/mod => lenient; high => no lenient),*

Black people would've been **3.3 pp less likely** to get lenient bail (91.5% vs. 94.8%) than white people

*(automating to recommendation generates a 65% smaller racial gap than observed)*

# Racial disparities in risk scores, recommendations, and outcomes

Concern about the algorithm distribution:
usage might widen racial disparities




After the recommendations implemented,

Black people were 9.3 pp **less likely** to get lenient bail (36.7% vs. 46%) than white people

*If bail automatically set by recommendations (low/mod => lenient; high => no lenient),*

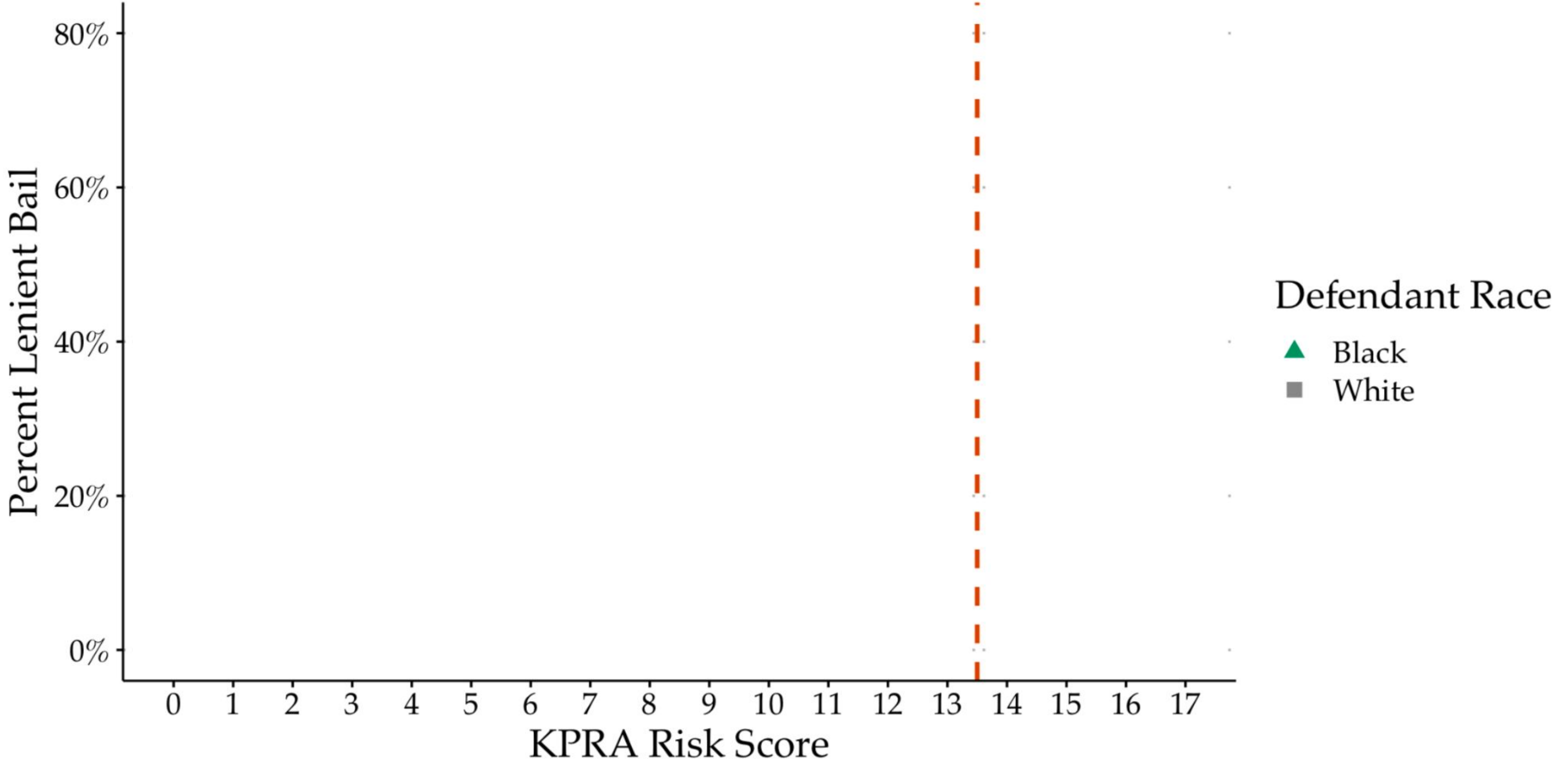
Black people would've been 3.3 pp **less likely** to get lenient bail (91.5% vs. 94.8%) than white people

*(automating to recommendation generates a 65% smaller racial gap than observed)*

***corollary: deviations from lenient recommendation vary by defendant race***

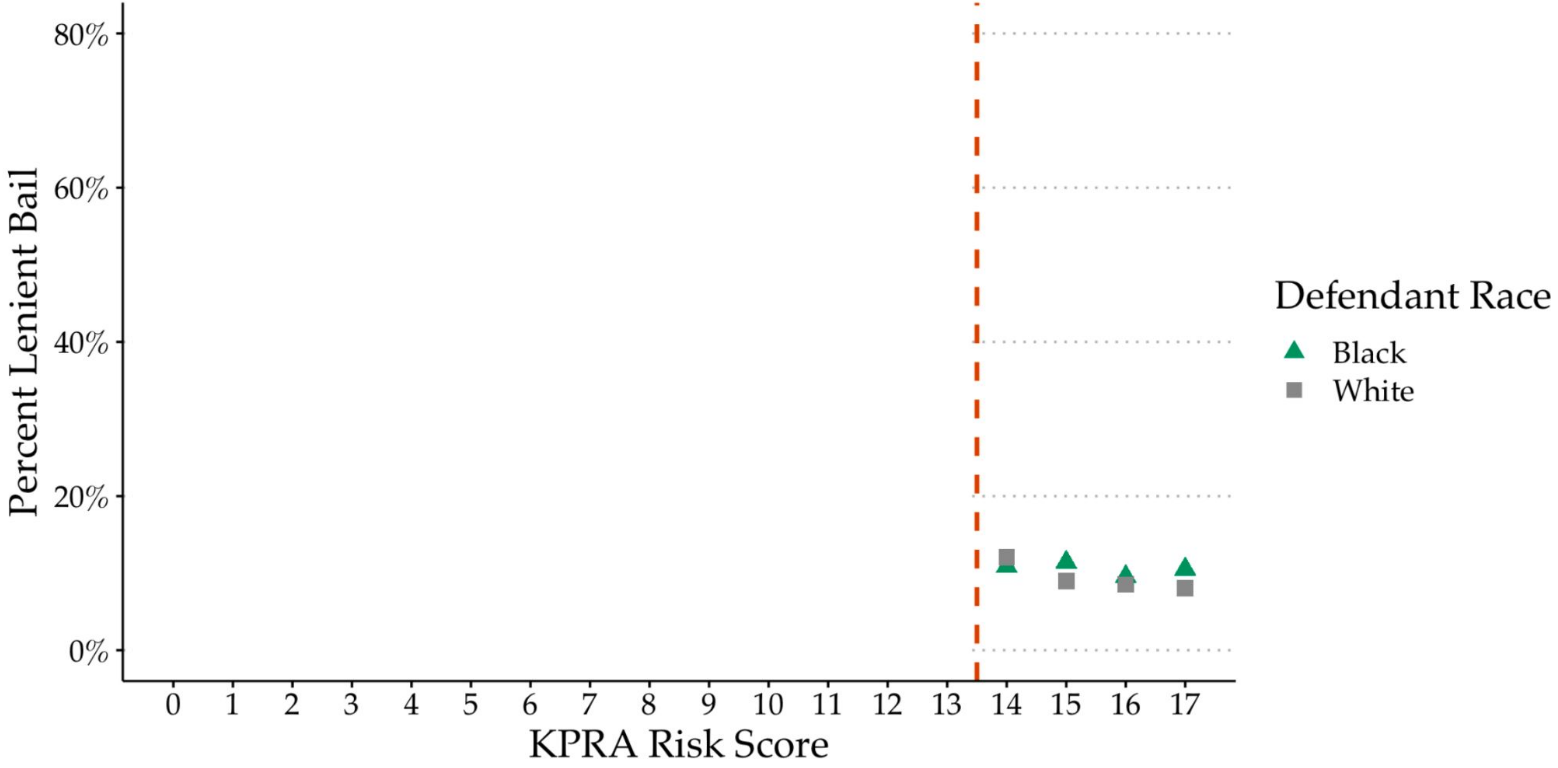
# Are differences in deviations due to different risk score distributions?

*Using data after algorithmic recommendations in place*



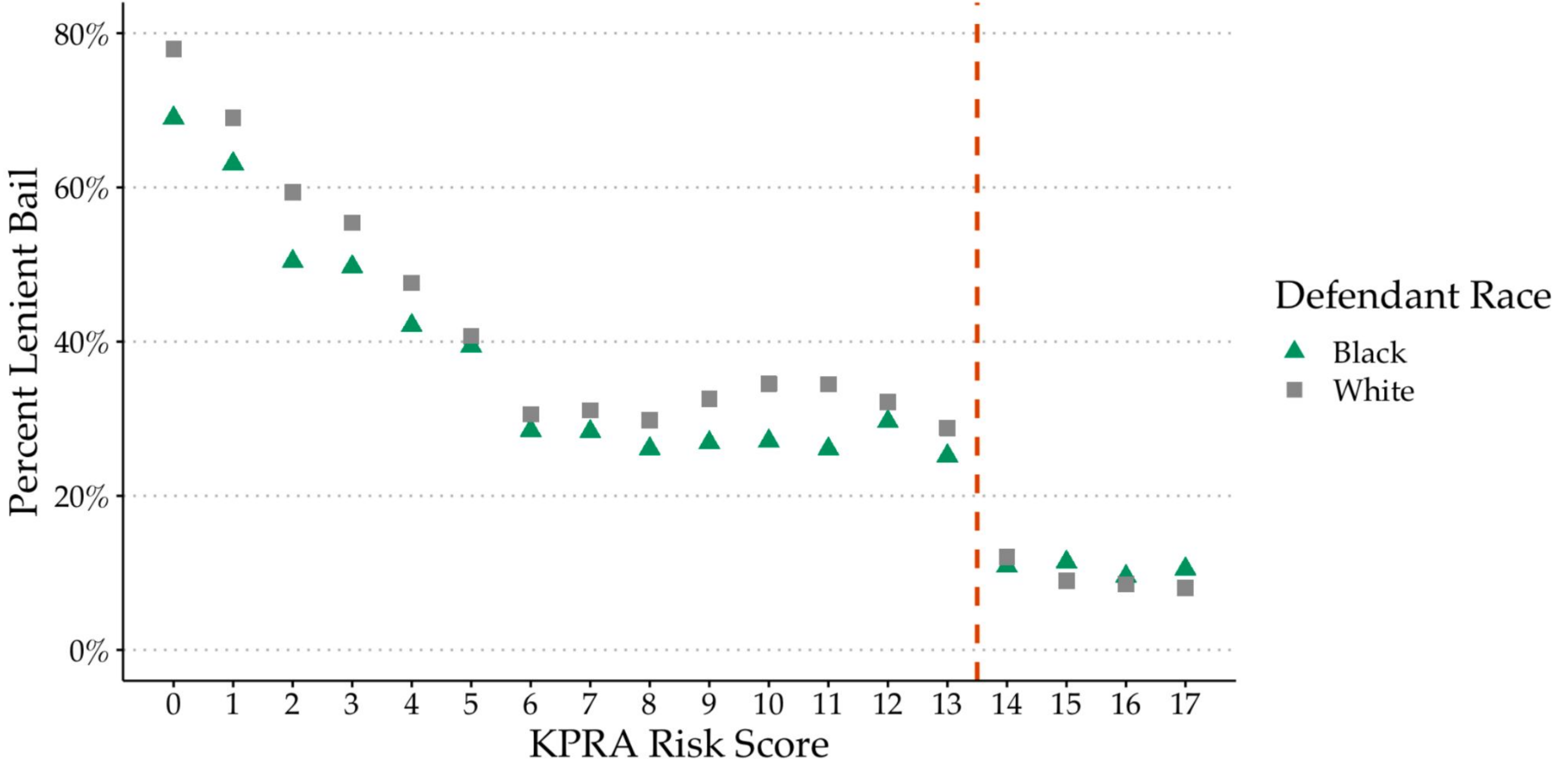
# Are differences in deviations due to different risk score distributions?

*Using data after algorithmic recommendations in place*



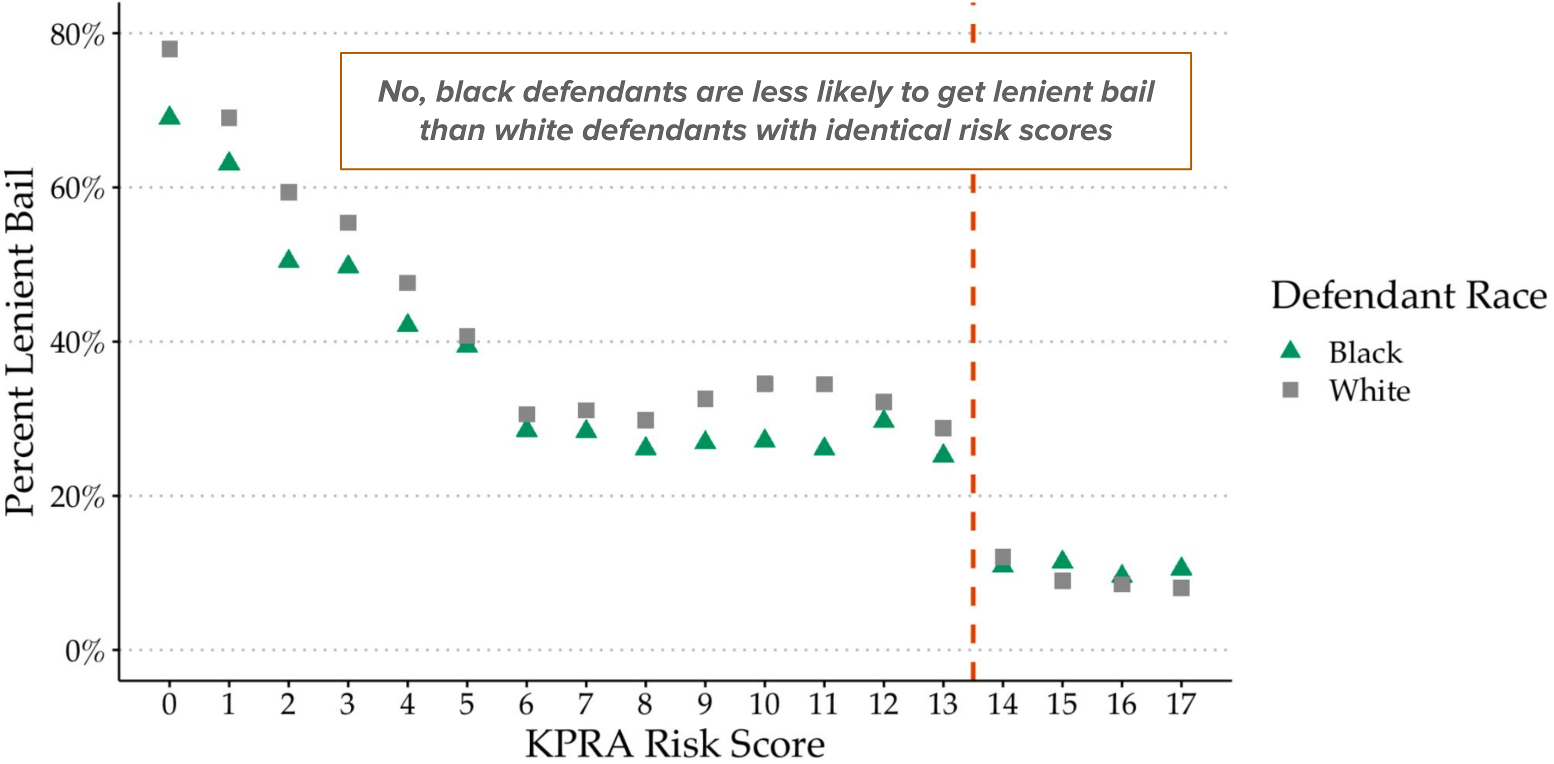
# Are differences in deviations due to different risk score distributions?

*Using data after algorithmic recommendations in place*



# Are differences in deviations due to different risk score distributions?

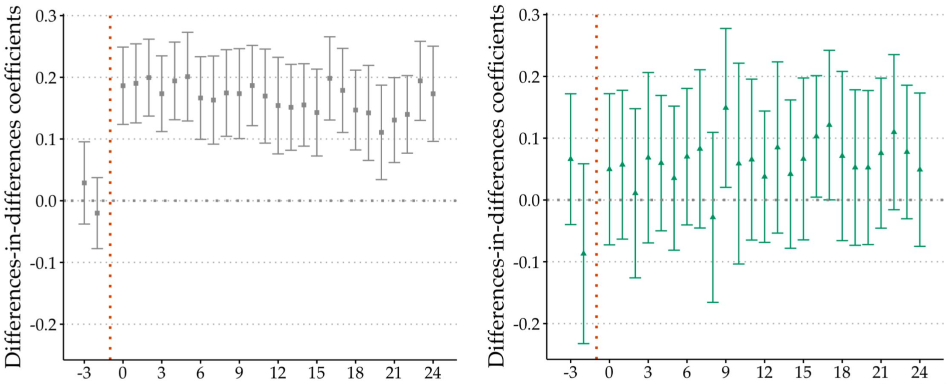
*Using data after algorithmic recommendations in place*




***No, black defendants are less likely to get lenient bail than white defendants with identical risk scores***

# Effects of algorithmic recommendation differ by defendant race

| | *Dependent variable: I(lenient bail)* | |
|---|---|---|
| | DD<br>*(White Defendants)* | DD<br>*(Black Defendants)* |
| | (1) | (2) |
| I(score<14) x Post | 0.168*** | 0.088*** |
| | (0.020) | (0.034) |
| Avg Dep Var (Pre-HB463) | 0.312 | 0.297 |



# Effects of algorithmic recommendation differ by defendant race

| | *Dependent variable: I(lenient bail)* | | |
|---|---|---|---|
| | DD *(White Defendants)* | DD *(Black Defendants)* | DDD |
| | (1) | (2) | (3) |
| I(score<14) x Post | 0.168*** | 0.088*** | 0.167*** |
| | (0.020) | (0.034) | (0.020) |
| I(score<14) x Black | | | 0.032 |
| | | | (0.029) |
| Post x Black | | | 0.009 |
| | | | (0.031) |
| I(score<14) x Post x Black | | | −0.083** |
| | | | (0.033) |
| Avg Dep Var (Pre-HB463) | 0.312 | 0.297 | 0.309 |

# What explains this heterogeneity?

| | Dependent variable: I(lenient bail) |
|---|---|
| | DDD |
| | (1) |
| I(score<14) x Post | 0.167*** |
| | (0.020) |
| I(score<14) x Black | 0.032 |
| | (0.030) |
| Post x Black | 0.009 |
| | (0.031) |
| I(score<14) x Post x Black | −0.083** |
| | (0.034) |
| Extra FEs | NA |

Are these differences
***within*** judges or ***between*** judges?

# What explains this heterogeneity?

| | *Dependent variable: I(lenient bail)* | | |
|---|---|---|---|
| | DDD | DDD | DDD |
| | (1) | (2) | (3) |
| I(score<14) x Post | 0.167*** | | |
| | (0.020) | | |
| I(score<14) x Black | 0.032 | ***Allow for time-score-varying judge FEs*** | ***Allow for time-score-varying county FEs*** |
| | (0.030) | | |
| Post x Black | 0.009 | | |
| | (0.031) | | |
| I(score<14) x Post x Black | −0.083** | | |
| | (0.034) | | |
| Extra FEs | NA | judge x under14 x post | county x under14 x post |

# What explains this heterogeneity?

| | *Dependent variable: I(lenient bail)* | | |
|---|---|---|---|
| | DDD | DDD | DDD |
| | (1) | (2) | (3) |
| I(score<14) x Post | 0.167*** | | |
| | (0.020) | | |
| I(score<14) x Black | 0.032 | −0.006 | −0.010 |
| | (0.030) | (0.035) | (0.028) |
| Post x Black | 0.009 | 0.006 | 0.008 |
| | (0.031) | (0.029) | (0.024) |
| I(score<14) x Post x Black | −0.083** | −0.020 | −0.026 |
| | (0.034) | (0.034) | (0.029) |
| Extra FEs | NA | judge x under14 x post | county x under14 x post |

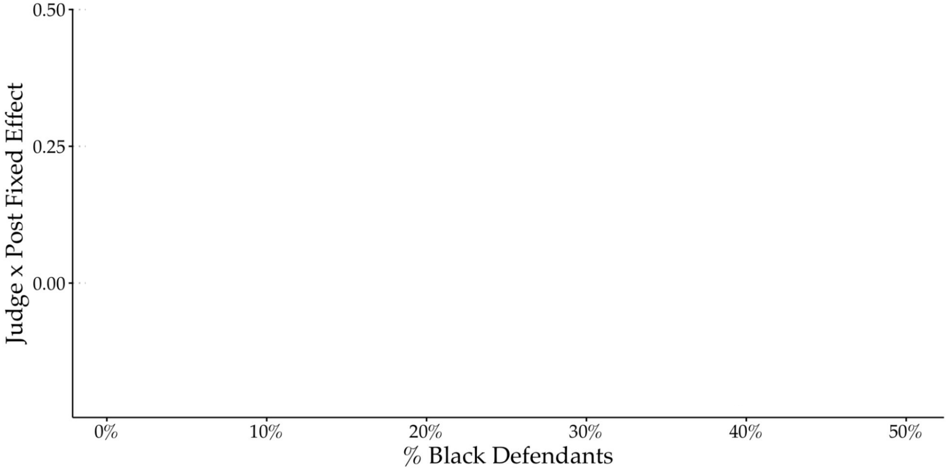
# Judges with more Black defendants respond less to lenient recommendations

| | *Dependent variable: I(lenient bail)* | | |
|---|---|---|---|
| | DDD | DDD | DDD |
| | (1) | (2) | (3) |
| I(score<14) x Post | 0.167*** | | |
| | (0.020) | | |
| I(score<14) x Black | 0.032 | −0.006 | −0.010 |
| | (0.030) | (0.035) | (0.028) |
| Post x Black | 0.009 | 0.006 | 0.008 |
| | (0.031) | (0.029) | (0.024) |
| I(score<14) x Post x Black | −0.083** | −0.020 | −0.026 |
| | (0.034) | (0.034) | (0.029) |
| Extra FEs | NA | judge x under14 x post | county x under14 x post |

# Judges with more Black defendants respond less to lenient recommendations

| | *Dependent variable: I(lenient bail)* | | |
|---|---|---|---|
| | DDD | DDD | DDD |
| | (1) | (2) | (3) |
| I(score<14) x Post | 0.167*** | | |
| | (0.020) | | |
| I(score<14) x Black | 0.032 | −0.006 | −0.010 |
| | (0.030) | (0.035) | (0.028) |
| Post x Black | 0.009 | 0.006 | 0.008 |
| | (0.031) | (0.029) | (0.024) |
| I(score<14) x Post x Black | −0.083** | −0.020 | −0.026 |
| | (0.034) | (0.034) | (0.029) |
| Extra FEs | NA | judge x under14 x post | county x under14 x post |

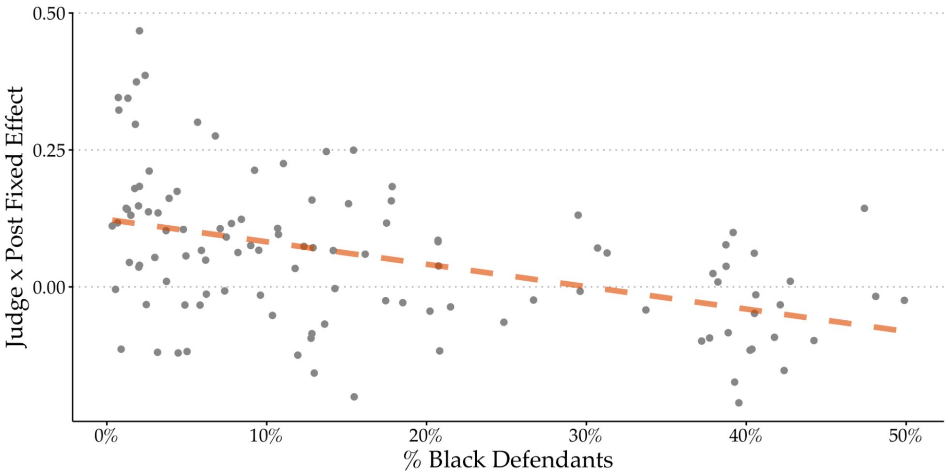
*Subset to cases with score <14,*
*Estimate judge x post FEs*
*(judge FE for post period)*



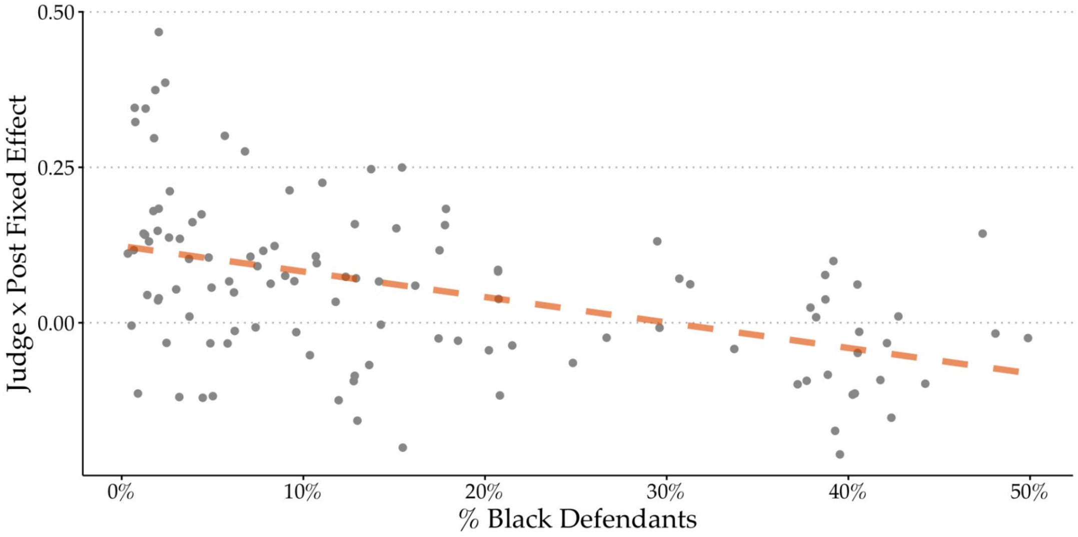
# Judges with more Black defendants respond less to lenient recommendations

| | *Dependent variable: I(lenient bail)* | | |
|---|---|---|---|
| | DDD | DDD | DDD |
| | (1) | (2) | (3) |
| I(score<14) x Post | 0.167*** | | |
| | (0.020) | | |
| I(score<14) x Black | 0.032 | −0.006 | −0.010 |
| | (0.030) | (0.035) | (0.028) |
| Post x Black | 0.009 | 0.006 | 0.008 |
| | (0.031) | (0.029) | (0.024) |
| I(score<14) x Post x Black | −0.083** | −0.020 | −0.026 |
| | (0.034) | (0.034) | (0.029) |
| Extra FEs | NA | judge x under14 x post | county x under14 x post |

*Subset to cases with score <14,*
*Estimate judge x post FEs*
*(judge FE for post period)*



# Why do they respond less?




Could this relationship be explained by...

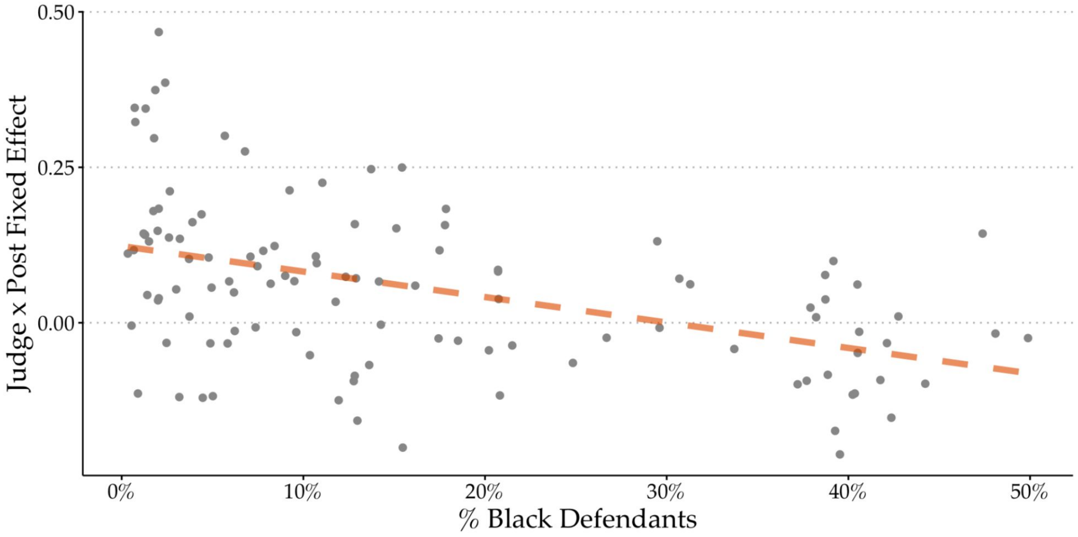
**Judge characteristics?**

- *Demographics (race, gender)*
- *Experience (years as judge)*
- *Election competitiveness*
- *Misconduct rates*

**County characteristics?**

- *Population*
- *Crime rates*

# Why do they respond less?




Could this relationship be explained by...

**Judge characteristics?**

- *Demographics (race, gender)*
- *Experience (years as judge)*
- *Election competitiveness*
- *Misconduct rates*

**County characteristics?**

- *Population*
- *Crime rates*

Data sources:

- **Judge demographics/experience:** hand-collect data from public profiles online, interviews with staff
- **Election competitiveness:** hand-collect data on 2010 local election PDFs
- **Misconduct rates:** calculate FTA/re-arrest rates by judge in pre-period
- **Population and crime rates:** county-level data from 2010 UCR data

# Why do they respond less?

| | *Dependent Variable = Judge x Post FE* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Black Defendants | $-0.374^{***}$ | | | | | |
| | (0.081) | | | | | |

**Judges who see 10 pp more Black defendants respond to the recommendation 3.7 pp less**

(25% drop from the 15 pp baseline effect)

# Why do they respond less?

| | Dependent Variable = Judge x Post FE | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Black Defendants | $-0.374^{***}$ | | | | | |
| | (0.081) | | | | | |
| | | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge |
| | | | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) |
| | | | | + FTA rate pre-<br>+ Rearrest rate pre- | + FTA rate pre-<br>+ Rearrest rate pre- | + FTA rate pre-<br>+ Rearrest rate pre- |
| | | | | | + County pop<br>+ Rural indicator | + County pop<br>+ Rural indicator |
| | | | | | | + Crime rate<br>+ Index crime rate<br>+ Prop crime rate<br>+ Violent crime rate |

**Judges who see 10 pp more Black defendants respond to the recommendation 3.7 pp less**

(25% drop from the 15 pp baseline effect)

# Other judge- and county-level covariates do not explain this

| | *Dependent Variable = Judge x Post FE* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Black Defendants | −0.374*** | −0.384*** | −0.377** | −0.323** | −0.307* | −0.374** |
| | (0.081) | (0.085) | (0.144) | (0.149) | (0.169) | (0.178) |
| | | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge |
| | | | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) |
| | | | | + FTA rate pre-<br>+ Rearrest rate pre- | + FTA rate pre-<br>+ Rearrest rate pre- | + FTA rate pre-<br>+ Rearrest rate pre- |
| | | | | | + County pop<br>+ Rural indicator | + County pop<br>+ Rural indicator |
| | | | | | | + Crime rate<br>+ Index crime rate<br>+ Prop crime rate<br>+ Violent crime rate |

**Judges who see 10 pp more Black defendants respond to the recommendation 3.7 pp less**

(25% drop from the 15 pp baseline effect)

# Other judge- and county-level covariates do not explain this

| | *Dependent Variable = Judge x Post FE* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Black Defendants | $-0.374^{***}$ | $-0.384^{***}$ | $-0.377^{**}$ | $-0.323^{**}$ | $-0.307^{*}$ | $-0.374^{**}$ |
| | (0.081) | (0.085) | (0.144) | (0.149) | (0.169) | (0.178) |
| | | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge | + Judge race<br>+ Judge gender<br>+ Years as judge |
| | | | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) | + Election contest<br>+ Contest in district<br>+ log(voters) |
| | | | | + FTA rate pre-<br>+ Rearrest rate pre- | + FTA rate pre-<br>+ Rearrest rate pre- | + FTA rate pre-<br>+ Rearrest rate pre- |
| | | | | | + County pop<br>+ Rural indicator | + County pop<br>+ Rural indicator |
| | | | | | | + Crime rate<br>+ Index crime rate<br>+ Prop crime rate<br>+ Violent crime rate |

**Judges who see 10 pp more Black defendants respond to the recommendation 3.7 pp less**

(25% drop from the 15 pp baseline effect)

Suggestive evidence:
Reputational cover recommendations provide depends on county demographics

*similar to Feigenberg + Miller (2021) finding of higher CJS severity in more racially heterogeneous places*

# Wrap-up

# Wrap-up

- Algorithmic recommendations are common practice in decision-making settings
- Effects of algorithmic recommendations are often hidden in study of effects of algorithms

This paper:

- ***Algorithmic recommendations have independent, economically meaningful effects***
- Why? ***Recommendations can change costs of mistakes***
- Deviations from recommendations can complicate effects on group inequality
    - ***Discretion matters even though algorithmic systems aim to limit its importance***

\end{talk}

# Thanks for coming!

You can email me here:

alex@albrightalex.com

*More info:*

*albrightalex.com*