# Algorithms as Prosecutors

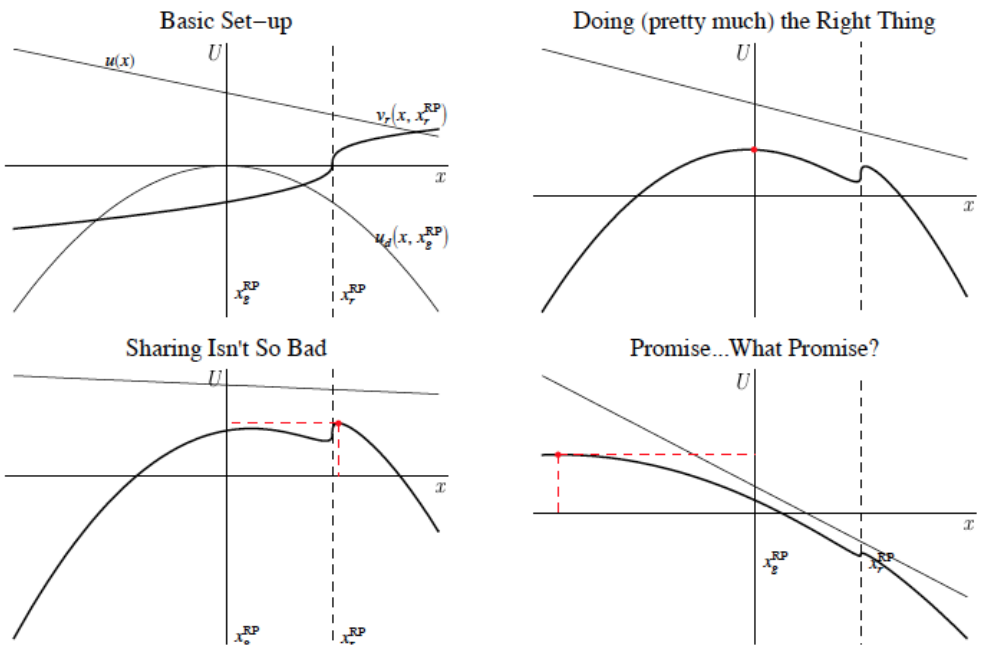
## Lowering Rearrest Rates Without Disparate Impacts and Identifying Defendant Characteristics 'Noisy' to Human Decision-Makers

D. Amaranto, E. Ash, D. L. Chen, L. Ren, C. Roper

# Legitimacy and Perceived Indifference

Justice: **equal treatment before the law** $(y = f(X) + \varepsilon, a \rightarrow X)$
**equality based on recognition of difference**
$(y \perp W, var(\varepsilon) \perp W, a \nrightarrow W)$

## Sympathy and Empathy

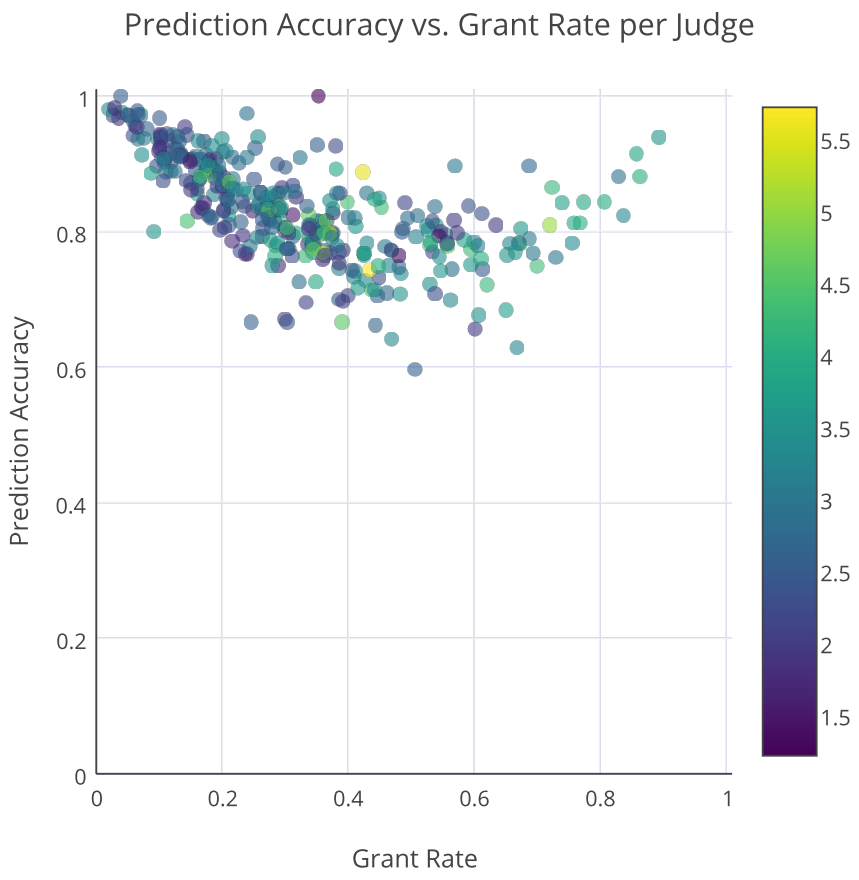



(Recognition-Respect theory)

"settings where people are closer to *indifference* among options are more likely to lead to detectable effects [of behavioral biases] outside of it." (Simonsohn 2011)

# Early Predictability

Significant inter-judge disparities in predictability raise questions of snap or predetermined judgement



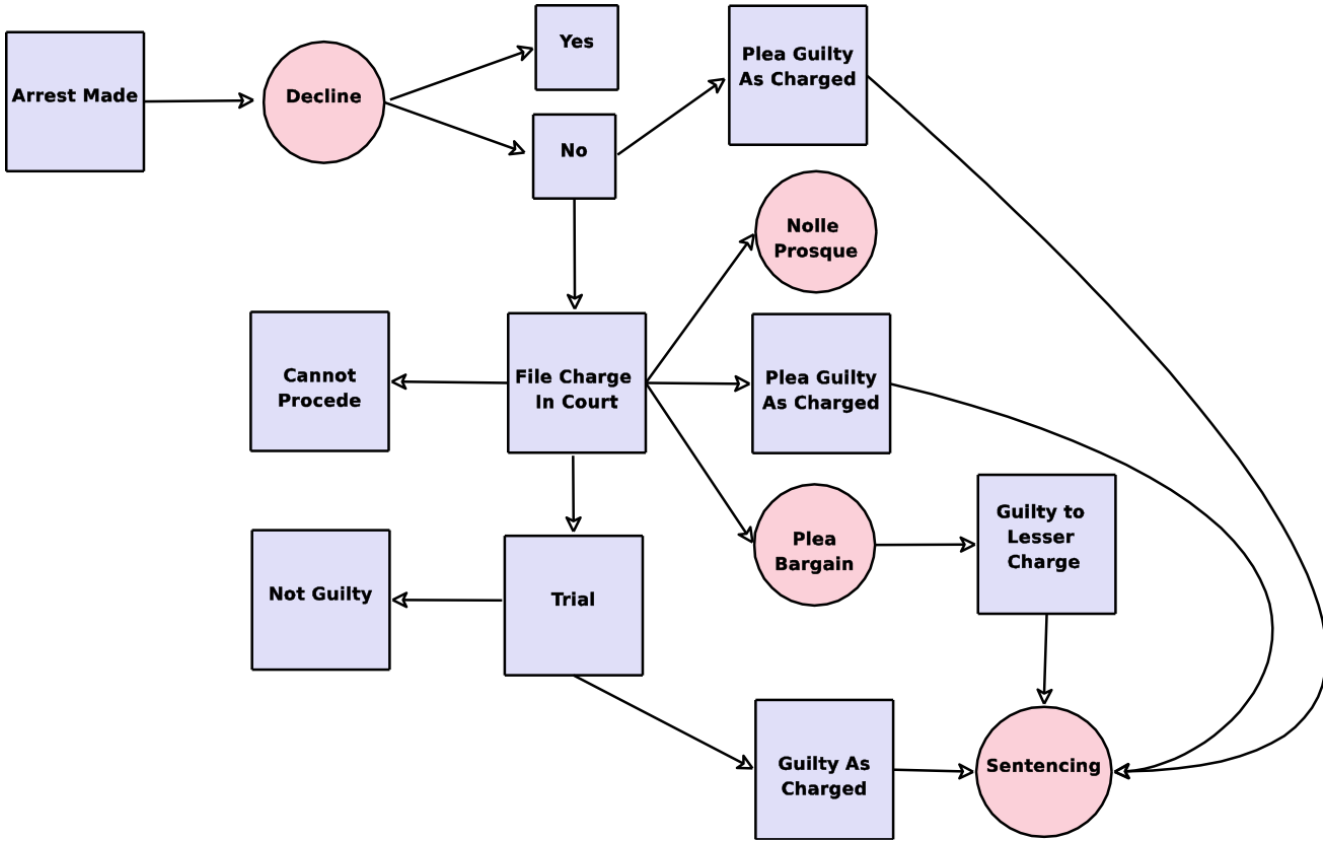

Judges with high grant rates use more hearing sessions (Chen, Dunn, Sagun, Sirin 2017 ICAIL)

- Information acquisition endogenous to preferences ("Redlining"; Brewer 1998)

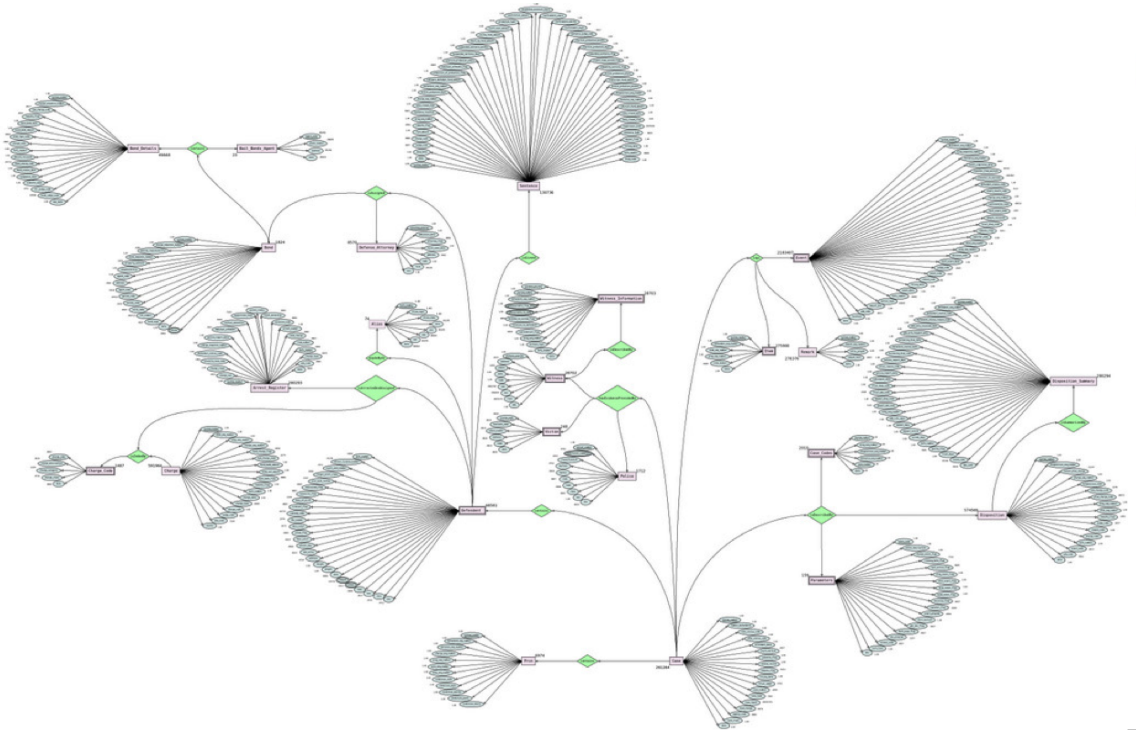
# Data

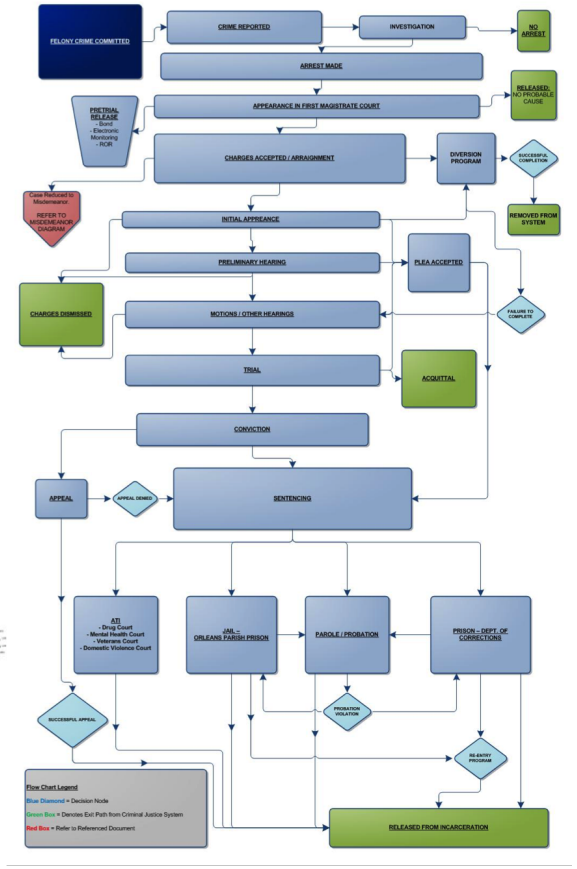
unique data from the New Orleans District Attorney's Office (1988-1999)

- detailed information regarding each individual offender and the corresponding prosecutor and judge
  - social security number (430K charges; 280K cases; 145K defendants)
  - victims, witnesses, defense attorneys, charges, police officers, sentencing, and many other topics (594 pg codebook)



# Original Dataset






- "skip tracer" to locate public demographic information on prosecutors and judges (including disciplinary actions, etc.) and their SSN
- some data on victims and witnesses
- 43% felony, optional; 17% felony, hard labor; 36% misdemeanor

# 'Data Generating Process'


**Figure 15.1.** Approximate Outcomes of 20,000,000 Felony Victimizations in the United States, 2007

- Other datasets are not linked
  - NCVS: victimization only
  - UCR/NBIRS: reports only
  - Fryer 2018: arrests context only
  - Random judge assignment studies: sentencing node only
  - NODA: arrests $\Longrightarrow$ sentence
    - ★ Prior studies examine (and can only examine) final decision node

# Broader Disparities in Criminal Justice

- Ferguson, Baltimore, Paris, Brussels, etc.
  - Motivations like trust in law / legitimacy of lawmaker hotly debated (Dworkin, Tyler, Acemoglu, Tirole, .. vs. Becker deterrence model)
  - Racial differences in police use of force (Fryer JPE 2018)


**Figure 15.1.** Approximate Outcomes of 20,000,000 Felony Victimizations in the United States, 2007

- Significant discretion in *whether* to charge a potential defendant (screening)
  - Interpretating sufficiency of police evidence
- Information about cases dropped by the prosecutor is unavailable
  - Prosecutors largely insulated from public accountability (Pfaff 2016)

# Broader Disparities in Criminal Justice

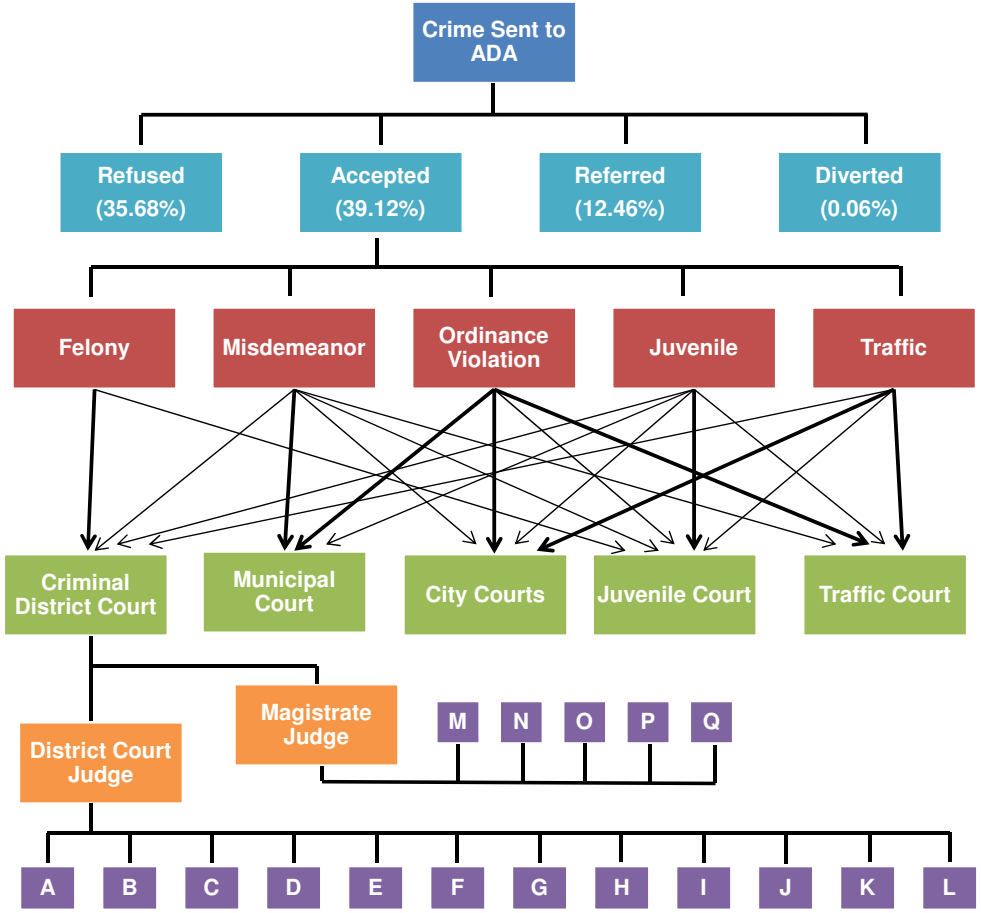
- Prosecutors very powerful
  - 1:15 — felony case a judge presides over in trial : the prosecutor decides the fate of 15 brought by police (eds. Wilson et al 2011)
  - From 1990-2010, roughly 50% of increase in felony filings comes from misdemeanors being charged as felonies (Pfaff 2016)
  - Controlling for *charge type* makes racial gap essentially disappear (!) (Rehavi and Starr JPE 2014)
- Decision *to charge* is equally important
  - Racial gap substantially magnifies when timeline pulled back

# Case Assignment

- "First, information on the case is received by a set of intake attorneys who routinely process the cases by collecting the potential defendant's rap sheet and other information. Then, once it has been processed, "[a] clerk then receives the file and assigns it to the screening attorney. The 'duty DA' for the day (**rotating duty**) handles everything that arrives on a given day except for the major crimes (such as homicide) assigned to a specialist. For special crimes, the clerk just assigns the case to whoever is next up **in the rotation** ([for example, a particular homicide prosecutor] receives two out of every five homicide cases, because there are two and a half homicide screeners)."

# Judge Assignment

- Once at the court, the cases were randomly assigned to a court section by the clerk's office (A-Q) public daily at noon.




- Felony cases must be scheduled randomly to prevent the district attorney from choosing a specific trial judge on the trial day and violating due process requirements. *State v. Simpson*, 551 So. 2d 1303 (La. 1989)

# Assessment of Random Assignment

- Check that judge characteristics are not correlated with defendant demographic (defendant race, sex, age, etc.) and other case characteristics.
- Judge leniency (a simplified Jackknife IV) and a collection of defendant traits. The judge leniency ($Z_{jt}$) is constructed as follows:

$$Z_{jt} = \frac{1}{n_{jt} - 1}\left(\sum_{k=1}^{n_{jt}} B_k - B_i\right) - \frac{1}{n_t - 1}\left(\sum_{k=1}^{n_t} B_k - B_i\right)$$

where $i$ denotes an individual case/charge, $j$ denotes the assigned judge, $t$ is the year of observation, $n_{jt}$ is the number of cases seen by a judge in year $t$, and $n_t$ is the number of cases seen by all judges in year $t$. For testing judge assignment, $B_i$ is a conviction decision.

# Assessment of Random Assignment

- Regression of judge harshness including case class by month-of-sentence fixed effects, judge-level clusters

| | (1) JudgeHarsh | (2) JudgeHarsh |
|---|---|---|
| criminal_flag | -0.00329 | 0.000548 |
| | (0.00760) | (0.00410) |
| dfdn_age | 0.000195 | 0.000213 |
| | (0.000132) | (0.000148) |
| dfdn_black_hair | 0.000882 | -0.0000624 |
| | (0.000977) | (0.000749) |
| dfdn_brown_skin | 0.00176 | 0.00235 |
| | (0.00209) | (0.00131) |
| dfdn_has_smt | 0.00229 | 0.00234 |
| | (0.00139) | (0.00141) |
| dfdn_height_feet | -0.00101 | 0.000136 |
| | (0.000888) | (0.000478) |
| dfdn_male | -0.000199 | 0.000525 |
| | (0.000801) | (0.000714) |
| dfdn_weight | 0.0000408 | 0.0000352* |
| | (0.0000229) | (0.0000162) |
| dfdn_white | 0.00132 | 0.000940 |
| | (0.00167) | (0.00136) |

Table: Testing for Random Assignment of Cases to Prosecutors

| **Dependent Variable:** | **Prosecutor Leniency** | | **Screened In (C** |
|---|---|---|---|
| **Pre-determined characteristics** | coef. | (s.e.) | coef. |
| Days from Police Report to Screening Date | -0.00959 | (5.691) | -2.186** |
| Days from Arrest to Screening Date | 5.243 | (13.80) | -16.95** |
| Predicted Screen | -0.000335 | (0.00230) | 0.104* |
| Missing Phone Number | -0.000826 | (0.00676) | -0.00530 |
| Days between Police Report and Arrest | 12.31 | (9.820) | -10.08** |
| Detained at End of Arrest Proceedings | 0.00205 | (0.00714) | 0.0284* |
| Detained at Start of Arrest Proceedings | 0.00432 | (0.00655) | 0.0333+ |
| Height | 0.172 | (0.107) | 0.134** |
| Male | 0.0217 | (0.0163) | 0.0267** |
| Weight | 1.071 | (1.108) | 0.435 |
| Birth Year | 1.473 | (1.010) | -0.550* |
| Out-of-state Drivers License | -0.00581 | (0.00441) | -0.00291* |
| Born out-of-state | -0.0110 | (0.00942) | 0.00848 |
| Race coded as Black | 0.000840 | (0.0121) | 0.00545 |
| Race coded as Asian | -0.00294+ | (0.00164) | -0.00135** |
| Race coded as Hispanic | -0.000329 | (0.00127) | 0.000428 |
| Race coded as Indian | 0.000169 | (0.000310) | -0.000190* |
| Race coded as Negro | 0.0270 | (0.0224) | 0.00658 |
| Race coded as Oriental | 0.0000244 | (0.00179) | -0.000274 |

Table: Testing for Random Assignment of Cases to Prosecutors

| **Dependent Variable:** | **Prosecutor Leniency** | | **Screened In (Case Accep** | |
|---|---|---|---|---|
| **Pre-determined characteristics** | coef. | (s.e.) | coef. | (s.e.) |
| Race coded as Other | -0.0000924 | (0.000983) | 0.000571 | (0.00044 |
| Race coded as White | -0.0247 | (0.0191) | -0.0112* | (0.0043 |
| Skin coded as Black | -0.00111 | (0.00164) | 0.000488 | (0.00058 |
| Skin coded as Brown | -0.00562 | (0.00598) | 0.00240 | (0.0017 |
| Skin coded as Dark | -0.000723 | (0.00239) | 0.000184 | (0.00035 |
| Skin coded as Fair | -0.00129 | (0.00489) | -0.00377** | (0.0012 |
| Skin coded as Light Brown | 0.00424 | (0.00825) | -0.000217 | (0.0021 |
| Skin coded as Light | -0.00130 | (0.000850) | 0.0000139 | (0.00020 |
| Skin coded as Medium | -0.00323 | (0.00241) | -0.000844+ | (0.00042 |
| Skin coded as Olive | -0.00297 | (0.00192) | -0.000239 | (0.00056 |
| Skin coded as Sallow | 0.000310* | (0.000144) | 0.0000619 | (0.00005 |
| Skin coded as Yellow | -0.000406 | (0.000355) | -0.000134 | (0.00012 |
| Skin coded as Dark Brown | 0.000285 | (0.0201) | 0.0173** | (0.0054 |
| Skin coded as Medium Brown | 0.0276+ | (0.0140) | -0.00741+ | (0.0042 |
| Skin coded as Ruddy | -0.0157 | (0.0173) | -0.00788* | (0.0036 |
| Eyes coded as Brown | 0.0104 | (0.0169) | 0.00728** | (0.0026 |
| Eyes coded as Blue | -0.0120 | (0.0113) | -0.00419** | (0.0015 |
| Eyes coded as Brown | -0.000320 | (0.00195) | -0.000934 | (0.00068 |
| Eyes coded as Green | -0.00671 | (0.00476) | -0.000324 | (0.00090 |

Table: Testing for Random Assignment of Cases to Prosecutors

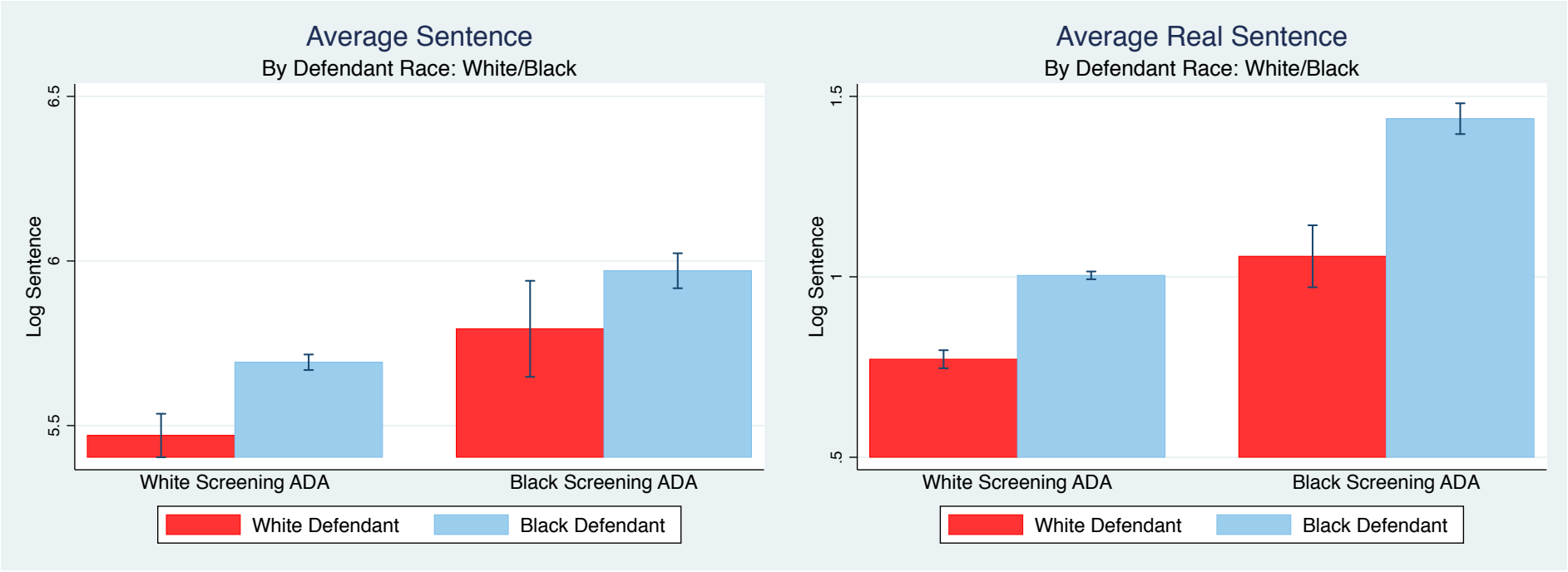
| **Dependent Variable:** | **Prosecutor Leniency** | | **Screened In (Case Accep** | |
|---|---|---|---|---|
| **Pre-determined characteristics** | coef. | (s.e.) | coef. | (s.e.) |
| Eyes coded as Hazel | 0.0105 | (0.00753) | -0.00187 | (0.00130 |
| Eyes coded as Grey | -0.00187+ | (0.000963) | 0.0000450 | (0.000205 |
| Hair coded as Brown | -0.00618 | (0.0156) | -0.00830* | (0.00414 |
| Hair coded as Black | 0.0235 | (0.0215) | 0.0129** | (0.00462 |
| Hair coded as Bald | -0.000223 | (0.00219) | -0.000232 | (0.000474 |
| Hair coded as Blond | -0.00534 | (0.00410) | -0.00252* | (0.00106 |
| Hair coded as Grey | -0.0135+ | (0.00800) | -0.00116 | (0.000870 |
| Hair coded as Red | 0.00220 | (0.00283) | -0.000744 | (0.000936 |
| Hair coded as Sandy | -0.000347+ | (0.000177) | 0.0000517 | (0.000070 |
| Hair coded as White | -0.0000129 | (0.000200) | 0.0000219 | (0.000076 |

(Features coded by police.)

# How Prosecutorial Discretion Affects Racial Disparities

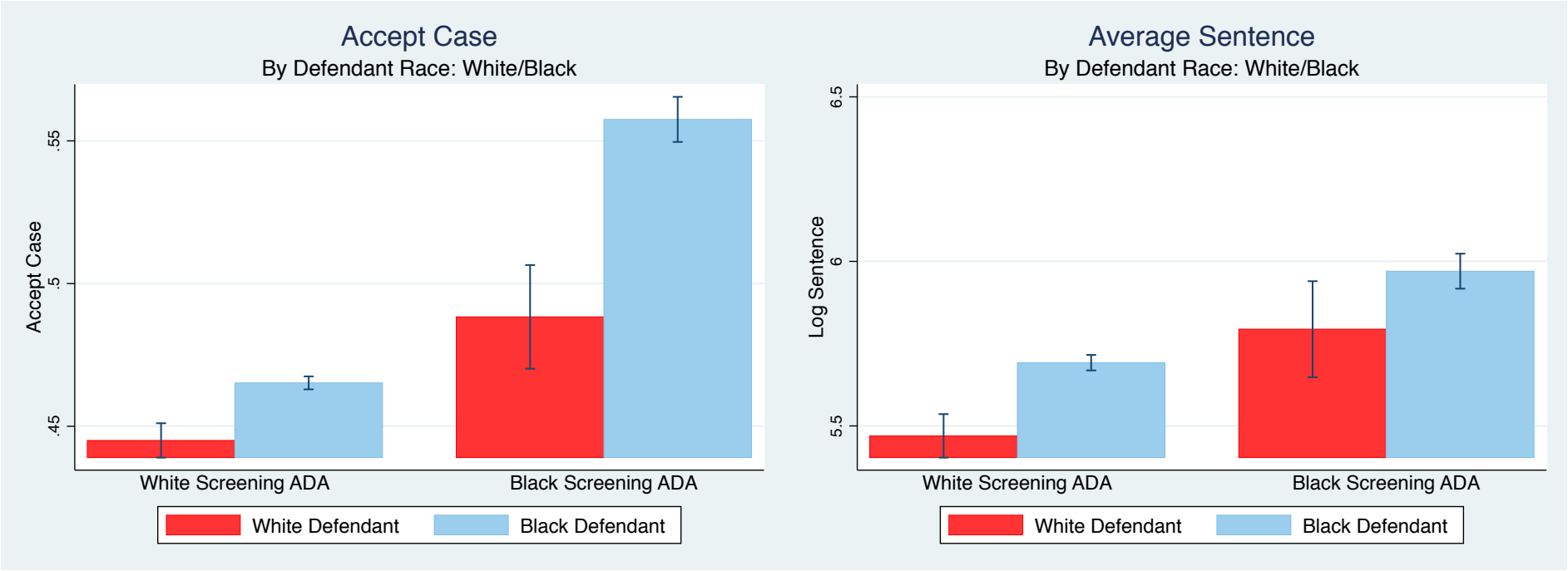
- Screening Magnifies Racial Sentencing Disparities
- Prosecutor Race Effects
- Racial Interactions in Courtrooms
- Algorithms as Prosecutors

# 1. Screening Increases Racial Gap



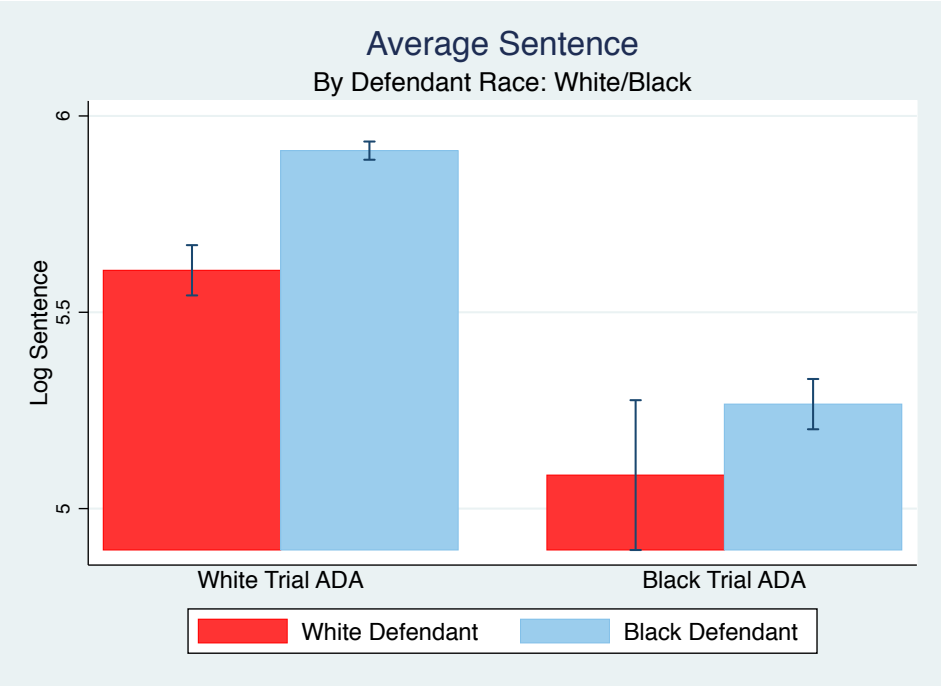

- Unconditional black-white sentence differences (on left)
- Since black defendants are less likely to be declined, racial disparity magnifies (on right)
  - Effects are quite large in log scale
  - Is statistical discrimination the reason for disparate screening?
    - ★ or does identity of decision-maker matter? (Anwar and Fang AER 2006)

# 2. White Screener Cases are Fewer and Leniently Sentenced



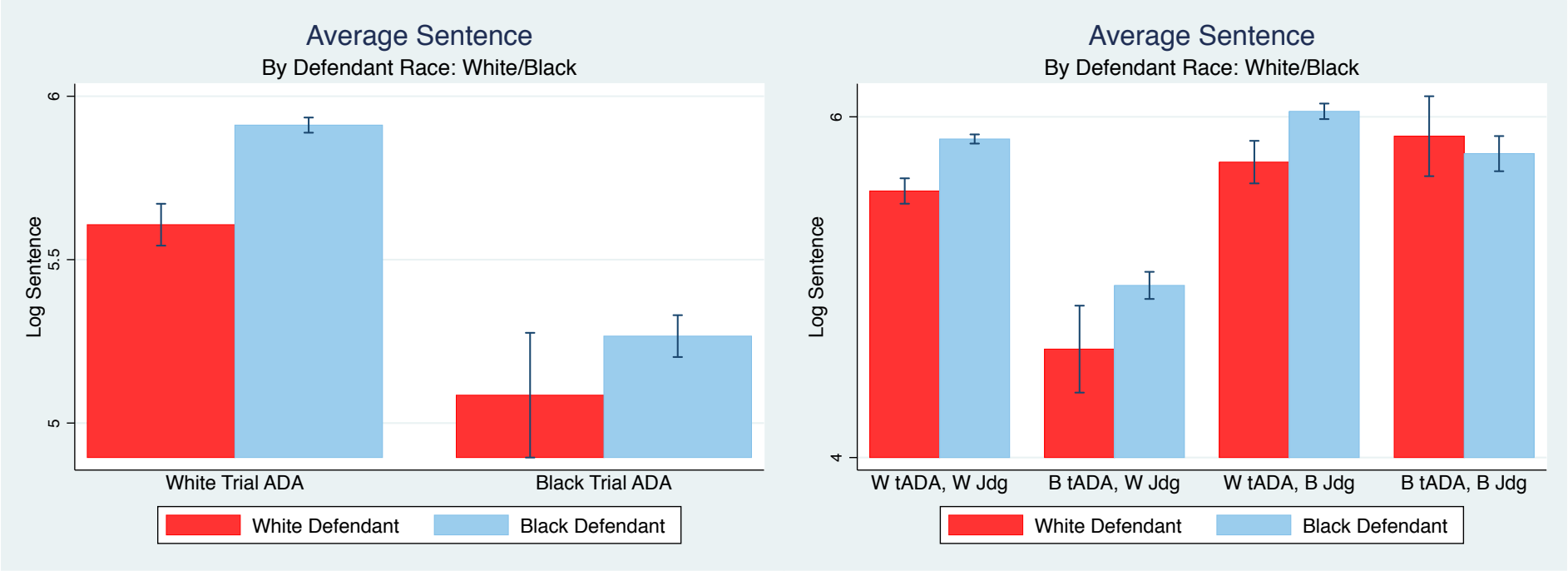

- Black defendants are screened in more (on left)
- White and black screeners let in different cases (on right)
  - If targeting the most severe ones, white screener cases should have *longer* sentences
  - Suggests not about statistical discrimination

# 3. White Trial Prosecutors Obtain Longer Sentences



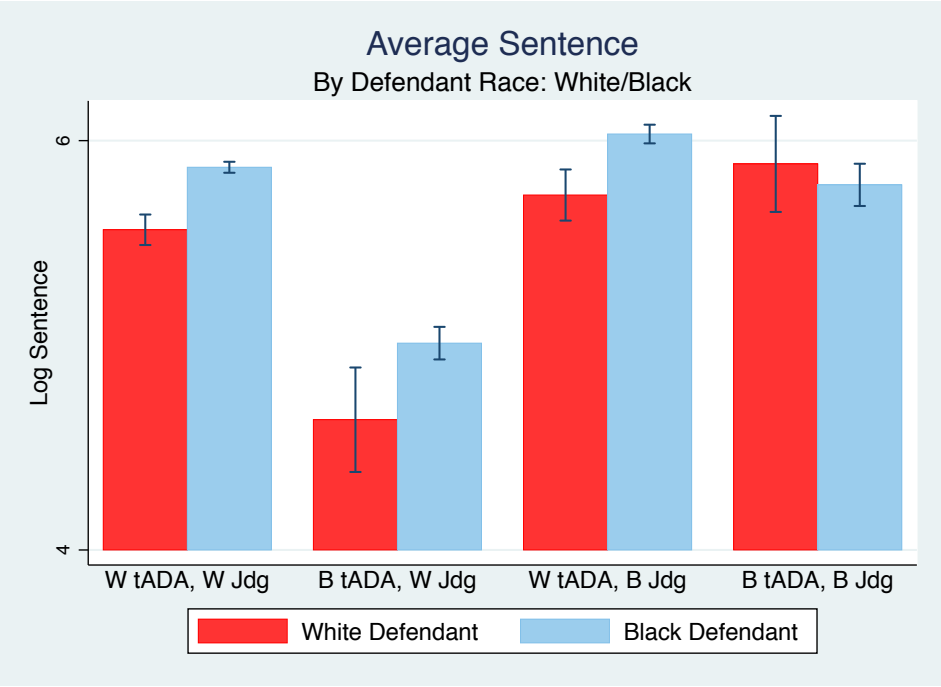

- Most District Attorneys are elected; want to appear tough-on-crime (Pfaff 2016)
- Why are white trial prosecutors more effective in this goal?

# 4. Black Trial Prosecutors + White Judges Render Shorter Sentences



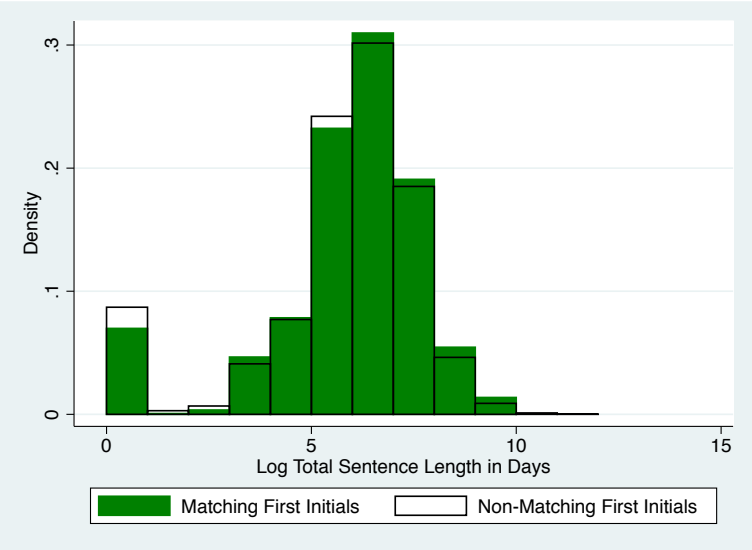

- The difference seems attributeable to the interaction of hierarchy and race
  - Black trial prosecutors + Black judges render similar average sentences as White trial prosecutors do
  - Effects are quite large in log scale (on right)

# 5. Black Trial Prosecutors + Black Judges Eliminate or Reverse Racial Sentencing Gap




- Hard to explain as statistical discrimination rather than ingroup bias
  - But ingroup bias by whom is not knowable without *benchmark*

# 6. Revealed Preference Indifference?




| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Log of Total Sentence in Days | | | | |
| Mean of dep. var. | 5.75 | | | | |
| First Letter Match | 0.102** | 0.0480 | 0.0308 | 0.0864 | 0.0889* |
| | (0.0442) | (0.0537) | (0.0930) | (0.0477) | (0.0498) |
| Defendants Sample: | Negro | Not Negro | Black | All | All |
| Judges Sample: | All | All | All | Black | White |
| Judge Fixed Effects | Y | Y | Y | Y | Y |
| Month-Sentence FE | Y | Y | Y | Y | Y |
| Case Class FE | Y | Y | Y | Y | Y |
| Observations | 33020 | 15840 | 10208 | 13441 | 35419 |
| R-squared | 0.443 | 0.492 | 0.539 | 0.457 | 0.462 |

- First letter name effects: 8% longer sentence lengths
  - conditional black-white sentence differences also $\sim$ 10% (Rehavi and Starr 2014)
- C1: effects are more salient for defendants classified (by the office) as Negroes
- C2, C3: effects are small and insignificant for those not classified as Negroes and for those classified as Blacks
  - "settings where people are closer to *indifference* among options are more likely to lead to detectable effects [of behavioral biases] outside of it." (Simonsohn 2011)

# 7. Long-Run Consequences of Screening? (w/ A. Philippe)

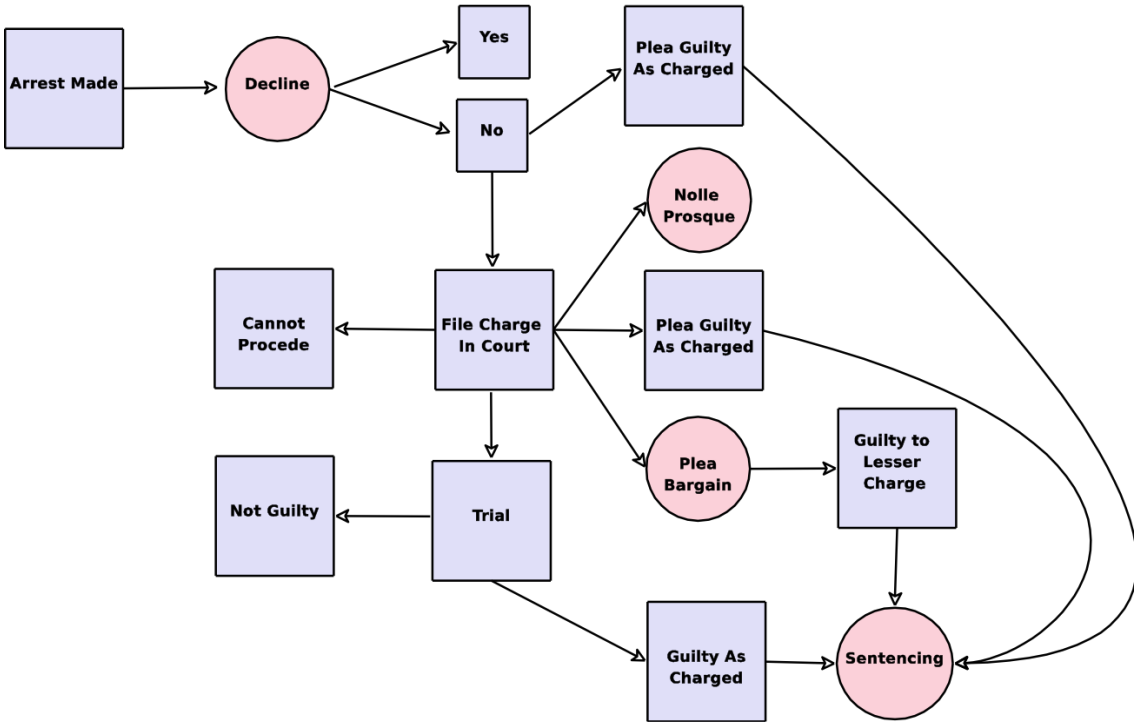

- Link to voter registries [preliminary]
  - Defendants assigned to harsher screeners
    - ★ more likely to register as Democrats
    - ★ more likely to vote
  - Ongoing analysis of families

# Signpost

- Prosecutors handle 15 times the cases that judges do (eds. Wilson et al 2011)
- We have seen that prosecutorial screening exacerbates racial disparities
  - The increase in disparities is not due to statistical discrimination
  - New evidence of
    - ★ revealed preference indifference,
    - ★ interaction of race and hierarchy, and
    - ★ ingroup bias
  - Screening may have long-run impacts

# Algorithms as Prosecutors?

- Experts have identified aggressiveness of prosecutors and plea bargaining as having particularly strong effects on US criminal justice system's ineffiencies & inequities. (Gopnik 2017; Miller and Wright 2002)
  - It has been argued that reduced emphasis on plea bargaining and heavier emphasis on screening could improve fairness & efficiency. (Miller and Wright 2002)
- We train a model of screening using rearrest as our target.
  - Assess performance compared to actual screeners. (Kleinberg et. al QJE 2018)
  - For a set charge rate, our model would reduce rearrest rates between 5-9%.
  - We use the model to understand how screeners select defendants to charge.



# Summary Statistics

| Variable | Mean | Std. Dev. |
|---|---|---|
| Binary Indicators | | |
| Criminal History | 0.8 | 0.4 |
| Detention at end | 0.04 | 0.21 |
| Habitual Offender | 0.05 | 0.21 |
| Detention at beginning | 0.05 | 0.21 |
| Juvenile | 0.25 | 0.43 |
| Male Prosecutor | 0.47 | 0.66 |
| Male Defendant | 0.81 | 0.42 |
| Multiple Defendants | 0.11 | 0.31 |
| Real-Valued Predictors | | |
| Total Number of Defendants | 1.18 | 0.72 |
| Days between police report and screening | 15.34 | 33.75 |
| Days between arrest and police report | 33.29 | 66.07 |

# Modeling

- One of the simplest ways to predict whether someone would be rearrested is a **decision tree** based on age and charge severity.
  - We constructed a baseline of max depth 4 using these two features.
  - 60% accuracy on the validation set and an F-score of 65%.
- We use an **ensemble model - gradient boosted trees**. F-score 77%.

| screening | arrest | habitual offender |
|---|---|---|
| date | date | criminal history |
| days since arrest | charge class (1-8 severity) | race (e.g., 85% black) |
| days since police report | charge type (arrest, information, indictment) | prosecutor |
| detention | law enforcement agency receiving credit | sex |
| beginning of arrest proceedings | total number of defendants | age |
| end of arrest proceedings | defendant | race (e.g., 79% white) |
| days between arrest & police report | juvenile, sex | days since bar admission |

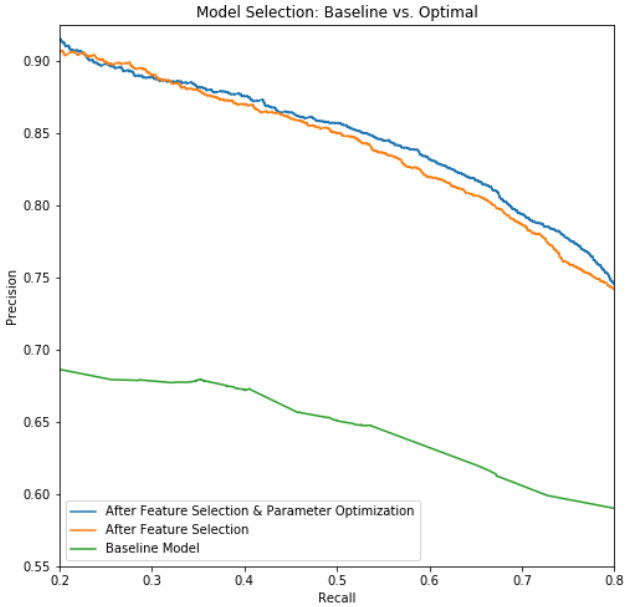
- Optimize hyperparameters - # trees, learning rate, depth, min. split

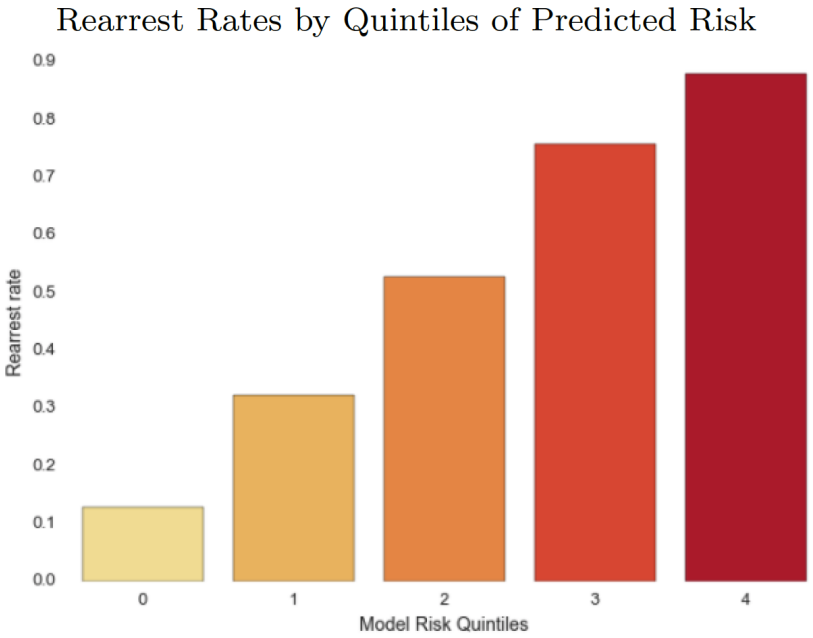
# Parameter Optimization

- All the models we trained use 50% as the probability threshold for each class, since the classes are balanced in our data (52% of arrestees were rearrested).
- For the gradient boosted trees, we trained the model with the following values of hyperparameters for a total of 81 models:
  - 1. number of estimators (100, 300, or 500)
  - 2. learning rate (.05, .1, .5)
  - 3. max depth (3, 5, 10)
  - 4. minimum samples split (2,4,8)
- All the gradient-boosted tree models use deviance as the loss function to be optimized and the Friedman mean squared error as the measure of the quality of a split. Our best-performing final gradient boosted trees model used 500 tree estimators, a learning rate of 0.05, a maximum depth of 5, and a minimum samples split threshold of 4.
- The final **confusion matrix** for the optimized model is below.

| | Prediction: Not Rearrested | Prediction: Rearrested |
|---|---|---|
| Not Rearrested | 3905 | 1391 |
| Rearrested | 1265 | 4433 |

# Model Performance




Gradient boosted trees substantially improves precision-recall over decision tree baseline

- Only considers cases where the arrestee was not charged. ◂ data cleaning
- The 64/16/20 split was stratified along the year of arrest so that the distribution of arrests over time was consistent among the training, validation, and test sets.
- To evaluate whether an arrestee was rearrested within a certain number of years, we truncated the data by that number of years from 1999, the last year of our data.
- The model also generates a probability of rearrest for an instance; a risk score. (Berk, ...)

Released arrestees predicted by our model to be risky were rearrested at a higher rate

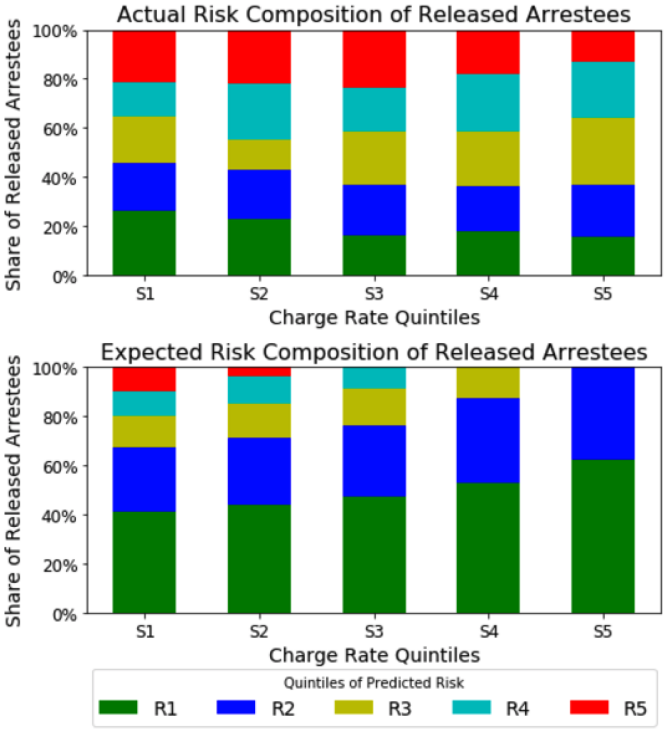
# Caveats

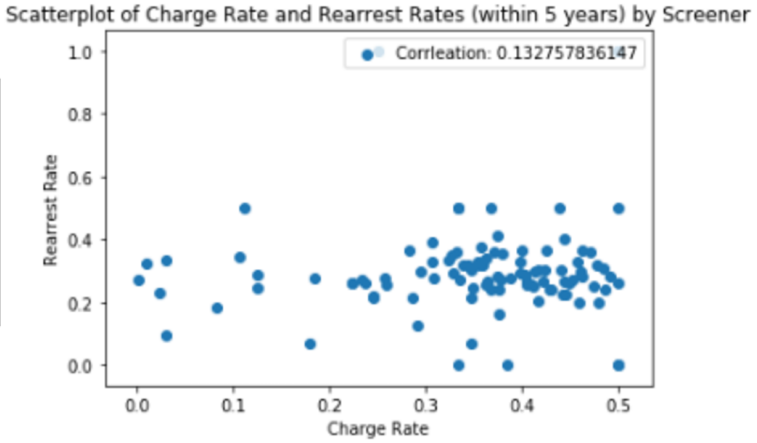
- Incomplete data
  - We were unable to track the arrest registry beyond 1999.
  - It is highly likely that at least some were rearrested in another district.
- Unobserved variables like educational background, socioeconomic status, and a host of other factors related to rearrest outcome.
  - Prosecutors observe things we do not, so would they do better?

# Using ML to Understand how Screeners Screen

How the screeners rank the risk of the arrestees is unobserved. However, **we can assess their implicit risk ranking by comparing the distribution of predicted risk of the arrestees charged by the "strict" and the "lenient" screeners.**

- Actual risk distribution amongst strict and lenient screeners differ from what we would expect to see if the screeners were releasing based on predicted risk.

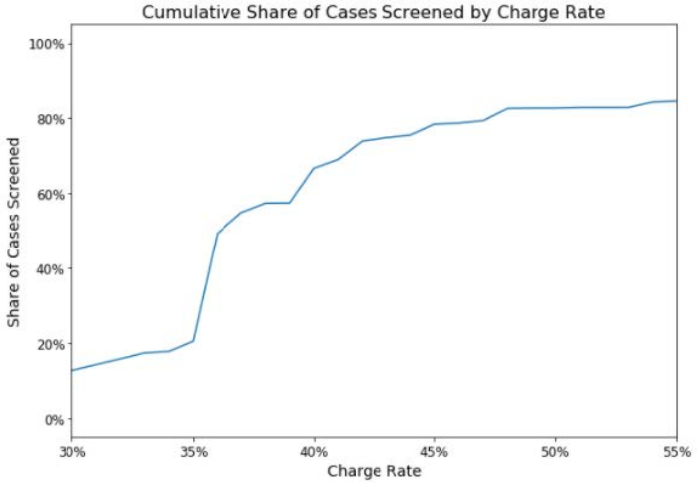


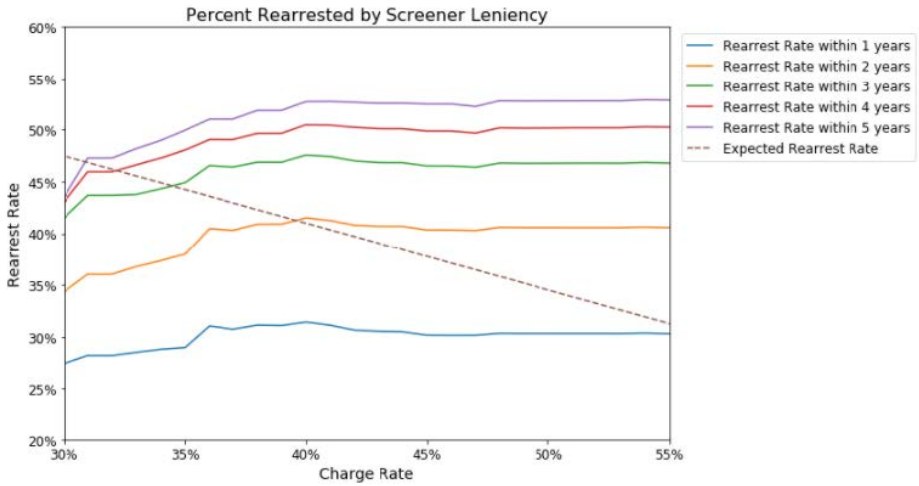




- Risk distribution of defendants released by screeners of increasing strictness from S1 to S5
- If screeners were to release defendants at random, we would expect to see an even distribution of predicted risk for each set of screeners – and even composition of R1 to R5

# Using ML to Understand how Screeners Screen

We can also assess the performance against actual rearrest rates.

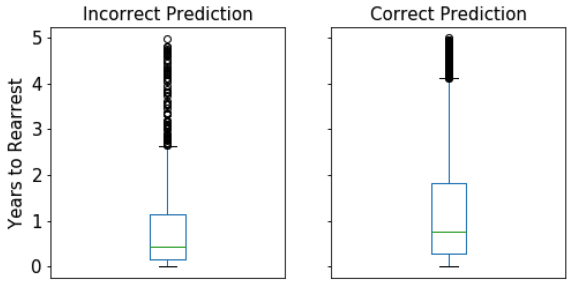


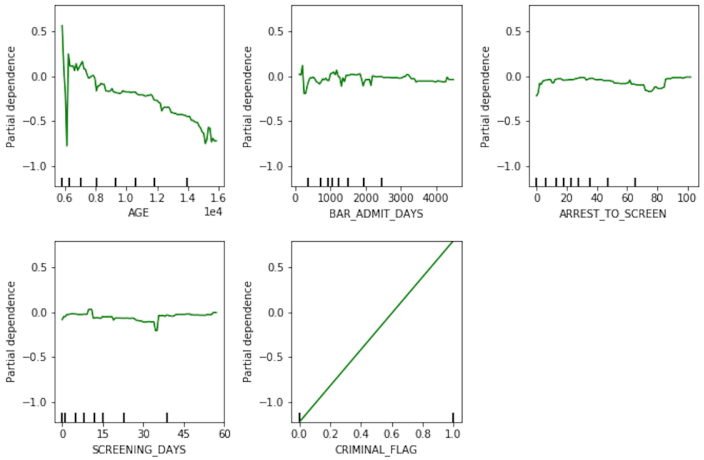

- We should observe a diagonal downwards slope from the upper left to the lower right if the screeners were releasing based on risk.
  - Instead, it is slightly *upward* sloping (more so, for 5 years out)

# Error Analysis and Time to Rearrest

- We better predict rearrest 5-years out than 1-year out

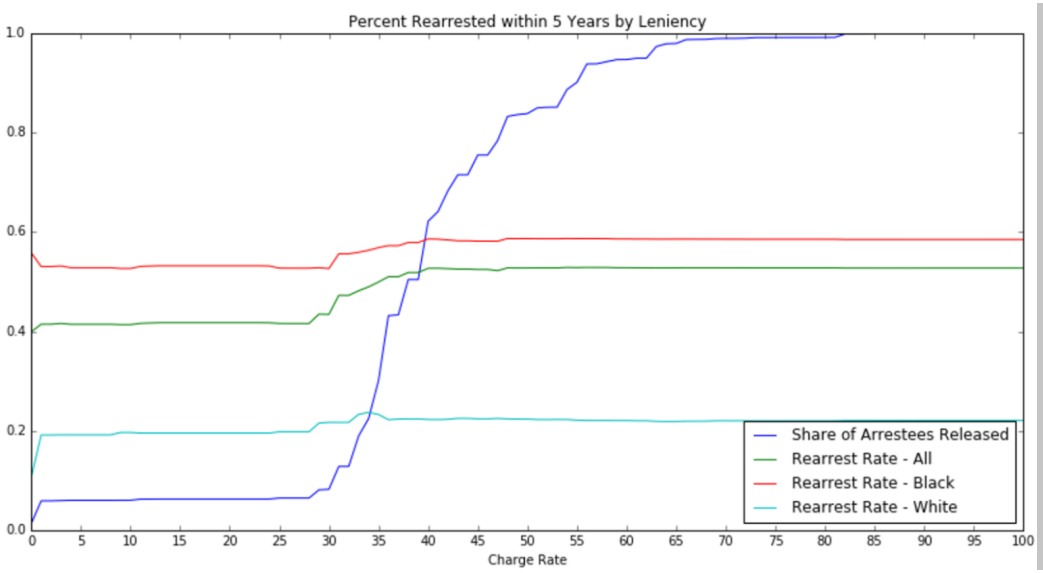


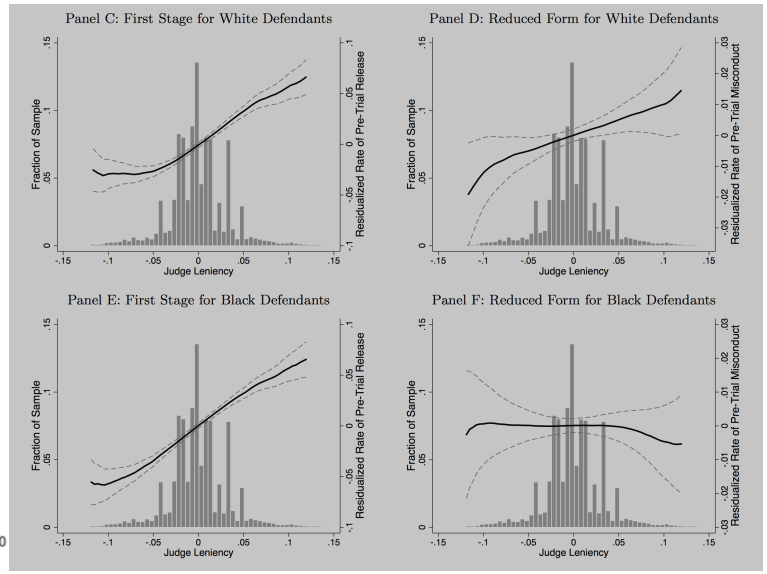

- Partial dependence plots of the top 5 "important" features.




- We removed features one at a time and observed how the F-score changed. Sex, age, and race were consistently among the features whose removal had the highest negative impact on performance.

# Using ML to Understand how Screeners Screen

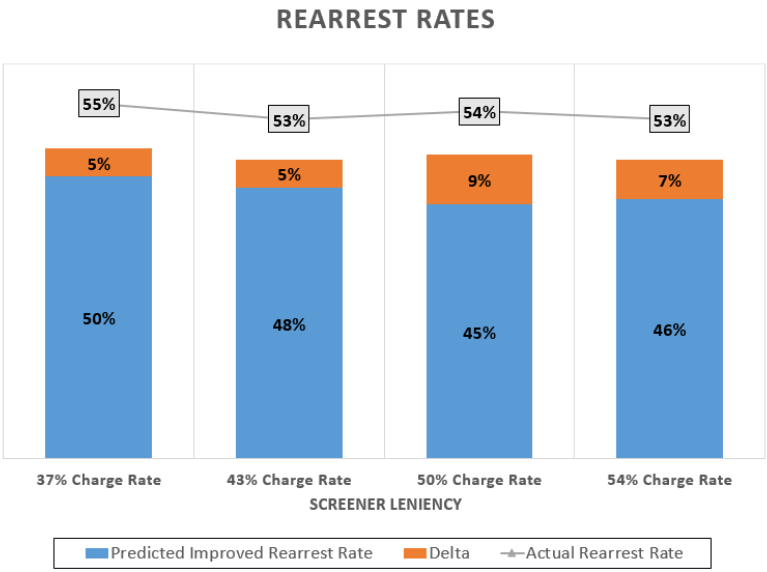



Actually, flat for Whites, *upward* slope for Blacks

- Judges released along "right" diagonal for Whites but not Blacks (Arnold et al QJE 2018)

# Potential Reduction in Rearrest from Using ML

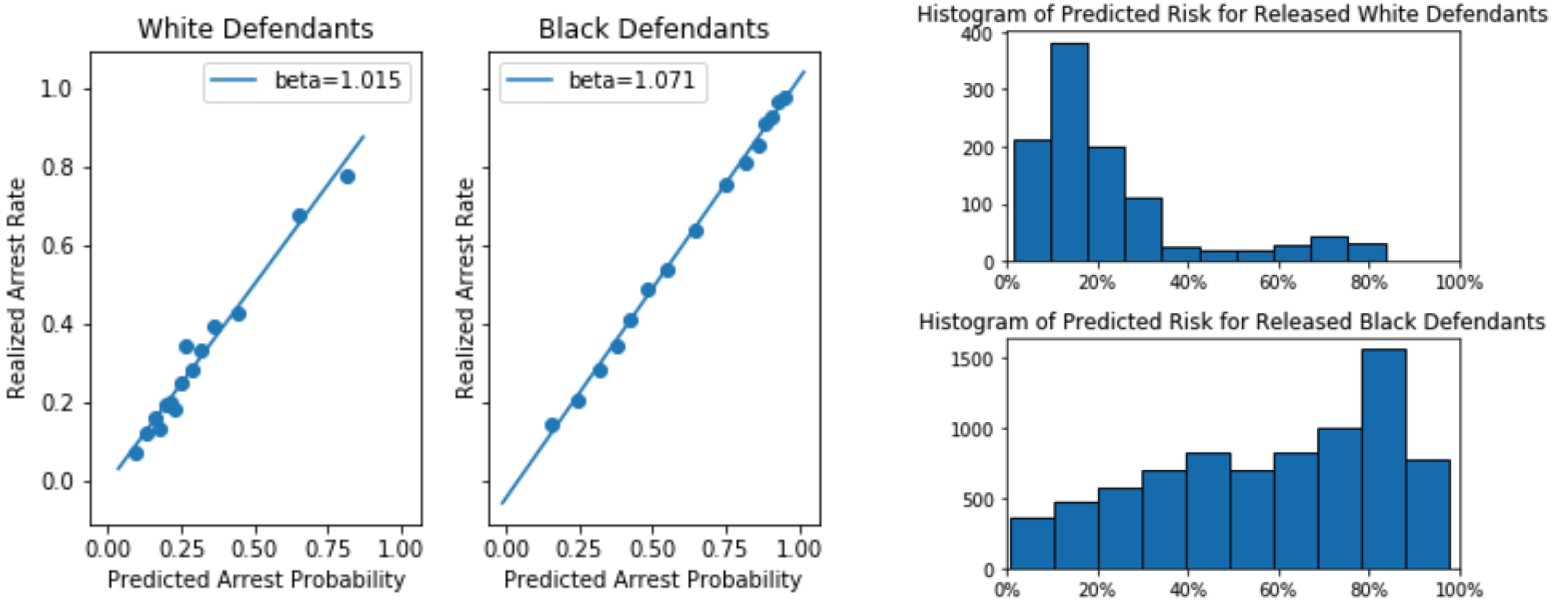



We analyze the "marginal" defendant.

- Given a screener(s), we define the marginal defendant as the defendant with the highest predicted risk that was seen and released by that screener(s).
  1. We calculate the additional number of arrestees that would need to be charged for the "lenient" group of screeners to reach the same charge rate as the next "strictest"
  2. We choose these "marginal" defendants based on estimated risk
  3. We assess actual rearrest rates of charging these additional "marginal" defendants
  4. If we arrive at a lower rearrest rate than the strict human screeners, then our model results in improvements in rearrest rates.
- **Inference**: We compared our results with the rearrest rates if we were to choose which the additional "marginal" defendants to charge at random (1000 times).

# Racial disparities did not increase with the model

- Consistent with "wrong" slope for Black defendants
- False positive rates were similar, 24% for whites, 21% for blacks.




- The algorithm tended to predict higher riskiness for black defendants.
  - Caveat: 11% of defendants are white and 86% of defendants are black.
  - Error rates not diagnostic for minimizing disparate impacts (S. Goel et al. 2018)

# Prosecutor Motivations

- It is possible that different predictions (e.g., maximizing *convictions* or *sentence length* or minimizing *time-to-trial*) are driving the decisions.
  - This does not, however, invalidate the comparison that we present. If a prosecutor's office cares about rearrest rates, and one would assume that they all do, then the results of a successful rearrest algorithm should still be relevant.
  - It's also possible these targets are associated for unobserved reasons.
- Could run separate models for other charge and rearrest types.
  - However, this would not affect the findings on equity:
    - ★ Blacks are screened in more within charge type
    - ★ Charge type is not a top-5 predictor of rearrest
    - ★ Should not observe the "wrong" slope (only for Blacks) even if unaware

# Conclusion

## Theory

- Revealed preference indifference, recognition-respect
  - rather than decision delay (Konovalov and Krajbich 2017)

## Method

- Evaluate when algorithms reduce disparate impacts

## Data

- Prosecutorial discretion contributes to unequal treatment
  - "the failure to punish everyday aggressions can be an important contributor to black disillusionment"

## Policy

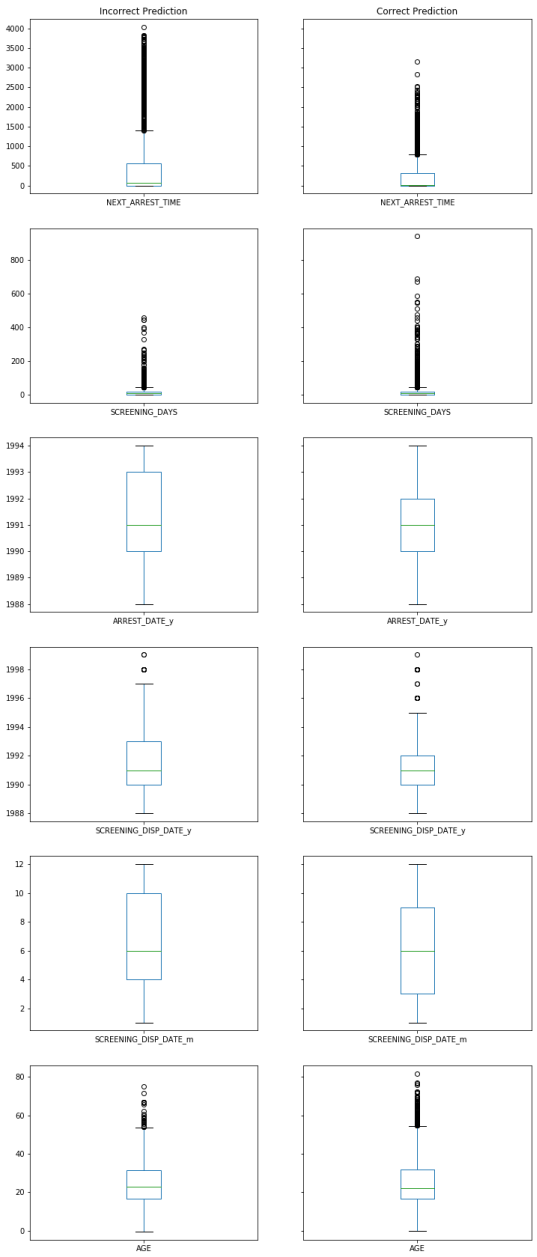
- Real-time decision support machine (historical caveat)
  - Lessen case burden with minimal impact on rearrest rates

## Next

- Screener traits and the characteristics of their declinations
  - Biases based on politics, age, experience, race, etc.
  - White screeners $\rightarrow \downarrow$sentences; what prosecutors maximize
  - Create a model to predict screeners and to compare the features predictive in screener vs. rearrest models.
    - ★ identify "mistakes"; personalized nudges to increase justice

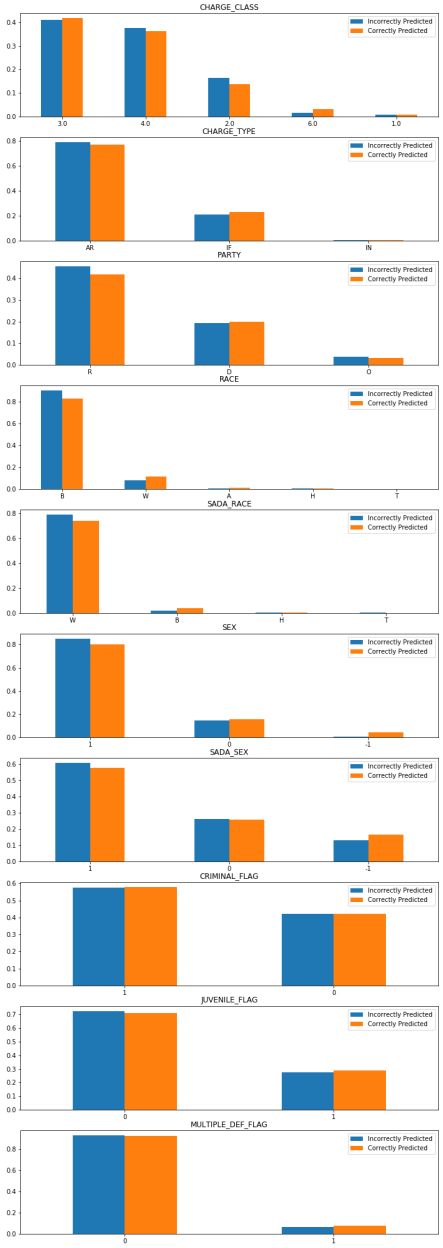
# Additional Material

# Other characteristics less related to accuracy



# Other characteristics less related to accuracy

Prosecutor Motivations



# Data Cleaning

- We imputed missing values for continuous-valued features based on the values of other non-missing features (i.e., imputing the value of the screening date from date of arrest and number of days between arrest and when the arrestee was seen by a screener).
  - For continuous-valued features where this was not possible, missing values were imputed as the mode. Using the codebook and lookup tables associated with the dataset, we flagged invalid values of binary and categorical features. Binary and categorical features were then transformed using one-hot encoding.
- We found that there were sometimes multiple arrests for a defendant on the same day. If charges for that defendant were **accepted for any** of the arrests on that day, we set the screening outcome code to indicate that charges were accepted.
- For example, if we wanted to create a target variable that indicates rearrest within 2 years, we would create that variable for defendants arrested between 1988 and 1997 so that we could conclusively determine whether an arrestee in 1997 was rearrested within two years. ◂ Model Performance

# Judge Assignment

- Felony cases must be scheduled randomly to prevent the district attorney from choosing a specific trial judge on the trial day and violating due process requirements. *State v. Simpson*, 551 So. 2d 1303 (La. 1989)
  - "A computer generated random allotment system" (La. Dist. Ct. R. 14.0, Appendix 14.0A) / a "bingo" system
  - "the allotment of cases shall be made **publicly** by classes **daily at noon** by the clerk or a deputy clerk selected by him, in the presence of the district attorney" (1991 La. R.S. 13:1343)
- Cases are classified into one of five classes
  - Random assignment without replacement
    - "Once a judge has been assigned a case from that class, he or she will not receive another assignment until all the other judges in that week's allotment have also received one case from that class."
  - At the start of each week, a small number of judges may be removed from the allotment process (based on vacation or other personal schedule issues)
    - "The eligible judges for the week's allotment determine how many marked balls go into the bingo machine."

# Assessment of Random Assignment

- Also assessed random assignment through simulation to compare the empirical distribution of case characteristics like race, sex, sentence length, incarceration rate, etc., to that found in simulated data.
  - Demographic composition of the judges and defendants and case characteristics of the defendants may change over time, and the simulation would reflect this.
  - Because of this variation in judge and defendant characteristics over time, it is necessary for the analysis to condition on short times when the random assignment of cases occurs.
  - A Monte Carlo simulation helps to overcome the finite sample bias because even though the overall sample is large, the sample observations are small within the short time periods that are of relevance.
- All observable case characteristic pairings of judge and defendants remain independent across all times, and we conclude that judges receive the same distribution of unobservable case characteristic such as criminal history, crime severity, etc., as well. ◂ Introduction

# Methods

- Log of 1+total sentence in days

$$SentenceLength_{ij} = F(t) + \alpha_1 \textbf{FirstInitialMatch}_{ij} + \epsilon_{ij} \tag{1}$$

- $\textbf{FirstInitialMatch}_{ij}$, is a dummy indicator for whether the first initial of the defendant and the judge match (6.4%)
- $F(t)$ can include a set of fixed effects for judge (our comparisons are made within-judge), month-of-sentence, case class, case class by month-of-sentence, charge code, and alphabetic identity of the letter
- Some specifications present placebo treatments:
  - whether the second letter of the name matches (15%)
  - whether the last letter of the name matches (10%)
  - randomly re-assigned first letter
- All specifications cluster standard errors at the judge level

# 6. Revealed Preference Indifference?

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Log of Total Sentence in Days | | | | | | | | | |
| Mean of dep. var. | 5.75 | | | | | | | | | |
| First Letter Match | 0.0799** | 0.0749* | 0.0853** | 0.0782** | 0.0686** | 0.0777** | 0.0754** | | | |
| | (0.0388) | (0.0383) | (0.0366) | (0.0362) | (0.0315) | (0.0350) | (0.0349) | | | |
| Second Letter Match | | | | | | | | -0.00494 | | |
| | | | | | | | | (0.0262) | | |
| Last Letter Match | | | | | | | | | 0.0334 | |
| | | | | | | | | | (0.0278) | |
| Resampled First Letter Match | | | | | | | | | | 0.0109 |
| | | | | | | | | | | (0.0256) |
| Judge Fixed Effects | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Month-Sentence FE | | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Case Class FE | | | Y | Y | Y | Y | Y | Y | Y | Y |
| Case x Month-Sent FE | | | | Y | | | | | | |
| Charge Code FE | | | | | Y | | | | | |
| Identity of Letter | | | | | | Y | | Y | Y | Y |
| Name Frequency | | | | | | | Y | | | |
| Observations | 48988 | 48988 | 48860 | 48860 | 48860 | 48860 | 48860 | 48859 | 48860 | 48860 |
| R-squared | 0.303 | 0.316 | 0.457 | 0.470 | 0.505 | 0.458 | 0.458 | 0.458 | 0.458 | 0.457 |

# 6. Revealed Preference Indifference?

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Log of Total Sentence in Days | | | | |
| Mean of dep. var. | 5.75 | 6.27 | 5.57 | 5.74 | 5.72 |
| First Letter Match | 0.0853** | 0.0454* | 0.0865** | 0.0811** | 0.0741* |
| | (0.0366) | (0.0260) | (0.0358) | (0.0358) | (0.0388) |
| Sample Restriction | None | $> 0$ | None | None | $< 8$ |
| Winsorize | None | None | 1% | 5% | None |
| Judge Fixed Effects | Y | Y | Y | Y | Y |
| Month-Sentence FE | Y | Y | Y | Y | Y |
| Case Class FE | Y | Y | Y | Y | Y |
| Observations | 48860 | 44775 | 48860 | 48860 | 46057 |
| R-squared | 0.457 | 0.497 | 0.458 | 0.456 | 0.436 |

- Robust to Outliers
  - C1: baseline
  - C2: drops sentences of length 0
  - C3, C4: winsorize at the 1% and 5% level
  - C5: restricts to sentences whose log length is less than 8

# Recap

**High** stakes decision making, if not for the DM, certainly for defendants

- more salient effects when a defendant is classified as "Negro"
  - impact of emotional shocks on judicial decisions affected minority defendants more (Eren et al. 2016)

**If** DM more susceptible to behavioral biases when more indifferent to the decision, then highly-trained professionals may also be susceptible to behavioral biases in other situations when indifferent

**Causal** effects explain a small portion of the overall variation

**(but** the behavioral response to perceived indifference by a DM is a different question altogether) ◂ Revealed Preference Indifference

# Random Assignment

- Assessed as before (jackknife leniency & Monte Carlo tests)
- "First, information on the case is received by a set of intake attorneys who routinely process the cases by collecting the potential defendant's rap sheet and other information. Then, once it has been processed, "[a] clerk then receives the file and assigns it to the screening attorney. The 'duty DA' for the day (**rotating duty**) handles everything that arrives on a given day except for the major crimes (such as homicide) assigned to a specialist. For special crimes, the clerk just assigns the case to whoever is next up in the rotation ([for example, a particular homicide prosecutor] receives two out of every five homicide cases, because there are two and a half homicide screeners)."

# Conclusion

## Theory

- Implicit Egoism (self-image of moral-decision makers matters)
  - Use economic tools to assess empirical basis for psychological phenomenona being questioned
    - ★ priming, gambler's fallacy, name letter effects, duty motivations

## Data

- Recent attention to police brutality (Fryer 2016)
  - Prosecutorial discretion is unexamined and important contributor to unequal treatment
    - ★ "the failure to punish everyday aggressions can be an important contributor to black disillusionment"